UNIVERSITY OF LIVERPOOL

FINDING HIDDEN STRUCTURES IN FINANCIAL NETWORKS

Thesis submitted in accordance with the requirements of

the University of Liverpool for the degree of Doctor of Philosophy

in Mathematical Sciences by

Stavros K. Stavroglou

30th September 2019

Table of Contents

Abstract

In this thesis we delve into the dynamic evolution of financial networks seeking real world meaning imprinted in their hidden structures. The hypothesis that permeates our research is that despite the stochastic behavior of single assets, when studied collectively there should be some emergent and persistent patterns which signal highly important information for scientists and policymakers alike.

Notwithstanding the insights from industry-standard methods, the hidden nature of causality remains a puzzling yet critical notion for effective decision-making. Financial markets are characterized by fluctuating interdependencies that can give rise to emergent phenomena such as bubbles or crashes. Motivated by these uncertainties, we designed a novel causality framework based on symbolic dynamics that probes beneath the surface of abstract causality and unveils the nature of causal interactions. We named our framework "pattern causality". This novel algorithm allows for a distinction between positive and negative interdependencies as well as a hybrid form that we refer to as "dark causality". We benchmark this method on asset pairs and on a network of sovereign credit default swaps, where the dominant form of interaction is that of dark causality. Our results are critical to financial advisors who have a fiduciary duty to their clients and retail investors.

Further contemplating upon the operational laws and concepts from complex systems, we composed a second algorithm out of the pattern causality framework with the purpose of capturing important aspects and interactions beyond stock markets. In an abstract complex network, it is an enigmatic and inspiring challenge to predict the actual interdependencies that comprise the structure of such systems, be it financial markets, ecosystems, or even the human brain. Particularly considering that the vital interdependencies underlying disparate real-world phenomena might be persistently hidden, the task of creating one algorithm to tackle them all seems daunting. Yet, our second algorithm is excellent at detecting the latent and elusive structures of complex systems. Our treatment utilizes short-term predictions from information embedded in reconstructed state space. Using a broad class of real-world applications, we are able to demonstrate our method's power to reconstruct the backbone of complex systems and simultaneously highlight their most fundamental operations. This last algorithm can serve as a tool for decision-makers and policymakers alike, and the demonstrated effectiveness establishes its potential for capturing hidden interactions in a much broader area of applications.

Acknowledgments

First and foremost, I am very grateful to my primary supervisor Dr. Athanasios Pantelous both for his scientific insights throughout my PhD as well as for securing a very gracious grant for the completion of our project. After four years, I consider Athanasios both a valuable mentor and a real friend. As a scientist he is characterised for his astute and practical thinking in our research directions. As a supervisor he has been very open-minded and embracing with regards to radical ideas, a very rare attribute for an academic. His attitude of inclusiveness, along with his spirited character, I find was a catalyst for our excellent collaboration. A life lesson I have adopted from him is his mantra for "zero-stress".

Next, I would like to thank my secondary supervisor Dr. Konstantin Zuev for his valuable input with regards to the mathematics of our algorithmic developments. His deep mathematical knowledge as well as his friendly demeanour were of utmost importance for the completion of this project.

Furthermore, I am very honored to have enjoyed the valuable mentoring of Professor H. Eugene Stanley who guided me in cementing the novel framework presented in this thesis. His time-tested and reputable acumen was inspirational for us to architect a framework which has the potential to usher a novel and solid discipline.

I am also indebted to my childhood friend Dimitri Pandeleakis, who helped me repeatedly with all aspects of programming. He sacrificed countless hours to both deliver end-results for my publications as well as teach me the nuts and bolts of software engineering.

Finally, I would like to thank my family for their omnipresent support and especially my brother George and my partner in life Lia for being always eager to discuss with me and challenge my ideas.

This project was funded by EPSRC and ESRC Centre for Doctoral Training in Quantification and Management of Risk and Uncertainty in Complex Systems and Environments (EP/L015927/1).

List of publications

1. Stavroglou, S. K., Pantelous, A. A., Soramaki, K., & Zuev, K. (2017). Causality networks of financial assets. *Journal of Network Theory in Finance*,
   DOI: https://doi.org/10.21314/JNTF.2017.029
   (This paper corresponds to **Chapter 2**)
2. Stavroglou, S. K., Pantelous, A. A., Stanley, H. E., & Zuev, K. M. (2019). Hidden interactions in financial markets. *Proceedings of the National Academy of Sciences of the United States of America*,
   DOI: https://doi.org/10.1073/pnas.1819449116
   (This paper corresponds to **Chapter 3**)
3. Stavroglou, S. K., Pantelous, A. A., Stanley, H. E., & Zuev, K. M. (2019). Unveiling Causal Interactions in Complex Systems. *Proceedings of the National Academy of Sciences of the United States of America*,
   DOI: https://doi.org/10.1073/pnas.1918269117
   (This paper corresponds to **Chapter 4**)

Illustrations

List of Tables

List of Figures

Notations

$LICC$	Linear intertemporal cross correlation
$NICC$	Nonlinear intertemporal cross correlation
LCo	Linear cointegration
NCo	Nonlinear cointegration
LGC	Linear Granger causality
NGC	Nonlinear Granger causality
SC	Shadow causality
HC	Hidden causality
Ω	A dynamical system evolving in an m-dimensional state space
M	The attractor of Ω consisting of all trajectories and possible states $\underline{\mu}(t)$ of the system. M is an E-dimensional attractor embedded in an m-dimensional state space ($E \leq m$). The state space contains the manifold and its dynamics and consists of the original E Cartesian coordinates (fundamental variables) of the system.
$\underline{\mu}(t)$	The point (vector) on M representing the state of the system at time t.
$X(t)$	State variables (time series) of the dynamical system Ω that operate as a function which maps points from Ω's attractor M to a real-valued scalar. X may correspond to Cartesian coordinates of the actual E-dimensional state space containing M.
t	Denotes time measured in discrete steps t_1, t_2, t_3, ... of X's temporal evolution.
L	The time series length, also called the library of the time series.
E	The embedding dimension of the attractor.
τ	The time lag we use to reconstruct a shadow attractor.
M_X	The shadow attractor reconstructed using time lags of $X(t)$.
$\underline{x}(t)$	The points (vectors) of M_X corresponding to the state of the system at time t.
h	The prediction horizon h steps ahead of current time t.
L_1	The Manhattan distance measured as $d_1\left(\underline{x}(t_1), \underline{x}(t_2)\right) = \sum_{i=1}^{E} \left\lvert \underline{x}\left(t_1^i\right) - \underline{x}\left(t_2^i\right) \right\rvert$

L_2	The Euclidean distance measured as $d_2\left(\underline{x}(t_1), \underline{x}(t_2)\right) = \sqrt{\sum_{i=1}^{E}(\underline{x}(t_1^i) - \underline{x}(t_2^i))^2}$
D_X	The distance matrix according to some metric (e.g. L_1or L_2).
$NN_{\underline{x}(t)}$	The nearest neighbors of $\underline{x}(t)$ according to D_X.
$S_{\underline{x}(t)}$	The vector of successive percentage changes of $\underline{x}(t)$.
$\hat{S}_{\underline{x}(t+h)}$	The vector of successive percentage changes of $\underline{x}(t+h)$.
$P_{\underline{x}(t)}$	The current pattern of $\underline{x}(t)$. For the trivial case of $E = 2$, we have three possible patterns: ↗, ➡, ↘.
$\hat{P}_{\underline{x}(t+h)}$	The estimated forecasted pattern of the affected variable, extracted as the signature of $\hat{S}_{\underline{x}(t+h)}$.
$PC[P_X, P_Y, t]$	The pattern causality (PC) matrix which is a 3D array with dimensions $(3^{E-1}, 3^{E-1}, L)$ which models the influence strength.
$P_X \parallel P_Y$	$\parallel$ denotes that the patterns are the same. See main diagonal (blue) of PC matrix.
$P_X \perp P_Y$	$\perp$ denotes that the patterns are opposites. See anti-diagonal (red) of PC matrix.
$P_X \nparallel P_Y$	$\nparallel$ denotes that the patterns are neither the same nor opposites. See all other elements (purple) of PC matrix.

1 Introduction

The course of nature and society can be tracked by regular small changes and abrupt monumental changes. The latter have the potential to veer the status quo.

In the modern economy abrupt changes usually take the form of speculative bubbles and stock market crashes. However, it would be naïve to assume that such events emerge all of a sudden in the economic system; rather they build up over time through the regular micro interactions of the actors and assets that comprise the system (i.e. traders, firms, governments and financial products). Aggregated over time, a regular interaction (e.g. speculation over exotic assets) can contribute to the formation of a strong interdependence (e.g. speculative bubble) and eventually may culminate into the emergence of catastrophic events (e.g. market crashes).

A case in point, at the dawn of the twenty-first century, we faced financial turbulence that rippled across stock markets throughout the world in a domino effect. The global financial crisis of 2007–9 still echoes in the minds of traders, scientists, and laymen; it shook the very foundations of traditional economics and standard forecasting methods, which failed not only to predict but also to prevent such a calamity.

One possible explanation of this failure could be the fact that traditional economics focuses on the isolated analysis of individual financial assets—e.g., market indices, bonds, stocks and commodities. “Isolated” in this instance means ignoring their interdependencies with other financial assets. Nonetheless, there are many reasons to defy the bias for isolated analysis.

First of all, institutional investors (who have increased their equity participation over the years according to Stambaugh, 2014, thus dominating the trading volume in equity markets) trade portfolios of at least a dozen assets which are subject to similar strategies. Even household investors, whose strategies tend to be significantly less sophisticated than their institutional counterparts (Frydman and Camerer, 2016), may trade more than a couple of assets in tandem. This common trading pattern suggests that the selling and buying orders of some bundles of assets are governed by the same person (or a small group of people) and are thus to some extent interrelated. An extraordinary but not rare case of such common trading patterns is the emergence of herding behavior, whence investors decision making gets synchronized (Jiang and Verardo, 2018). When a critical mass of investors demonstrates such herding behavior, the market usually starts behaving as a “hive” and as a result may be prone to irrational choices and often leads to crashes, like the bursting of the dot-com bubble.

Second, if we think of the asset prices as financial proxies of companies (equities, bonds), national governments (sovereign bonds), and essential market products (commodities), we cannot neglect the causal interactions between them. For example, an increase in oil prices can cause a decrease in demand for airline services; an innovation in the telecommunications industry can be patented and reverse the market share of competitors; an increase in the rates of government bonds directly affects stock market investments.

Finally, the psychology of people involved in a specific (national) stock market may be at least partly affected and in turn affect the fluctuations of stock markets in other countries—e.g., the news of a stock market crash can cause fear and panic beyond a nation's borders, affecting the international theater of transactions. The aforementioned herding behavior might thus lead to financial contagion, i.e. events in one market affecting markets in faraway regions. Add to these the cases of large capitalization international funds that invest in many stock markets, and we have more than enough reason to scrutinize the interdependencies of financial assets.

These interdependencies give rise to connected structures of assets that together comprise the ebb and flow of global markets or, hereafter, financial networks. Our motivation is to explore and scrutinize the dynamic evolution of such networks quantitatively. Our hypothesis is that even in the most stochastic setting of individual components, when studied collectively there should be some persistent patterns that cause grand scale phenomena. Such patterns could be, for example, a) specific stock prices predicting consistently other stocks' prices, b) increased synchronization (co-movement) of many stock prices before/during/after crashes, or c) feedback loops where one stock price influences another in a certain direction.

The greatest challenge strikes from the very conception of such thinking: there is no "right" way of establishing influence from one asset on another. For example, two asset prices could move together, implying correlation, but this could happen out of sheer coincidence as well. A surge in one asset price could cause a surge in another but this could be just a random incident. Therefore, the first research question that naturally arises from this hypothesis is: "*How can we establish whether X causes Y?*" For one, *X* and *Y* can refer to stock prices, but even more abstractly, to time-series in general. Correlation- and causality-inspired methods are the first that must be considered. However, in order to unveil the "true" structure of a network of interdependent entities, as accurately as possible, we need to be wary of mirage dependencies. Such dependencies can be inferred (mistakenly) for various reasons, such as using linear tools to discover an unknowingly nonlinear dependency.

Notwithstanding the imminent difficulty of finding the "true" hidden structures of financial networks, we begin by contemplating upon the primary methods used in the literature and understand them both at a theoretical level (how does each method infer a dependency from *X* to *Y*) and in an experimental setup that seeks out potential signals before the financial crisis of 2007-9. Overall, the methods we examine either infer relationships via a form of correlation or via a regression-based technique. Roughly speaking, these methods suggest that *X* influences *Y* if such a relationship is captured with a given measure for a pre-defined period of time. This free parameter of the length of time is actually a scourge for both scientists and practitioners. Stock price *X* might seem to influence stock price *Y* considering a rolling window of 10 days, but this relationship might as well disappear when considering a rolling window of 30 days. Thus, the first problem that arises from mainstream methods is the manual "choice of time", transforming the quest to answer the first research question into a need to actually invent a purposeful tool for that end.

Before creating a new tool, I had been tampering with those methods and a second research question emerged: "*In what way does X cause Y?*" Mainstream methods suggest that *X* can either cause *Y* in the same direction (positive influence) or in the opposite direction (negative influence). This rationale may make sense when we want a very rough estimation of an underlying relationship, but if we want a deeper understanding of how *X* causes *Y* then what about $Y = f(x) = \sin(X)$? In short term periods, *X* is causing *Y* either in the same direction or in the opposite direction. Yet overall *X* is causing fluctuations on *Y*, thus using correlations and regressions is ill-advised because these methods infer the type of influence dependent upon the pre-defined length of time we choose. This second problem, regarding the nature of influence, is even harder in the stock market where any stock price relationships are far from a pure sinusoidal function. Thus, the dual thinking of explicitly positive or negative does not inspire deeper scrutiny of interdependencies, thereby giving rise to the need for a more detailed framework.

To examine the initial hypothesis and analyze the collective behavior of assets while addressing the two research questions, the need for a new framework is unavoidable. However, creating a new framework for causality should both remove the free choice of time period as well as liberate the nature of influence from the strict dualism of positive and negative interdependencies. Completing such a task will deliver a new tool into the hands of scientists and practitioners for exploring and finding hidden structures in financial networks. What naturally arises then is a third research question: "*Does this new tool provide insights for*

complex networks beyond finance?" Answering this final question requires the study of networks beyond finance, such as ecosystems or brain functions.

In Chapter 2 we begin by studying the evolutionary behavior of the average causality in a given financial network with the purpose of answering the first research question regarding the basic ways we infer statistical influence. This analysis is conducted for eight causality methods. Our results ascertain the existence of significant interdependencies between assets throughout the time period examined. According to Hawawini (1980), Atchinson *et al.* (1987), Lo and MacKinlay (1990), Chowdhury (1991) and Fiedor (2014a, b) the very existence of causality in financial assets challenges the foundations of the well-known "efficient market hypothesis".

For our analysis, we consider that the global crisis started on August 9, 2007. More specifically, results from linear intertemporal cross-correlation do not imply any predictive capabilities, given that the average causality it measures begins to drop in parallel with the development of the global financial crisis, rather than before it. However, average causality as measured by nonlinear intertemporal cross-correlation could be a candidate for an early warning signal, given that almost half-a-year before the crash (at the outset of 2007) it started dropping to all-time lows, then started to rise more steeply than ever before as the crisis unfolded. Linear cointegration analysis does not show any different causality behavior before or during the global crisis. Its nonlinear counterpart exhibits a marked plunge just three months before the global crisis and then rises steeply. Both linear and nonlinear Granger causality showed no change in their average causality before or during the global crisis. Despite the fact that shadow causality (based on mutual information) displays no forecasting capabilities, it somehow "fits" the characteristics of the crisis period by displaying a dramatic drop right after the "birth" of the crisis after a protracted upward trend. Similarly, hidden causality (based on transfer entropy) simply changes its trend from horizontally fluctuating to slightly downward after the onset of the crisis. With the exception of nonlinear intertemporal cross-correlation and nonlinear cointegration, the methods simply follow the emergence of financial turbulence, serving at best as contemporaneous crisis indicators and not as early warning signals. However, these results by no means suggest that the remaining six causality methods are incapable of serving as early warning signals, only that they must be subjected to further scrutiny.

In Chapter 3, having tampered with the methods from Chapter 2, we introduce a novel framework for the identification and quantification of the nature of causality which we termed "pattern causality". Pattern causality is purposefully built to answer the second research question regarding the nature of causality while at the same time removing the free parameter

for the choice of time-period to consider. To achieve this, our method uses the whole time-period at once in order to attempt and predict *Y*'s next value based on *X*'s entire historical evolution up until now. This novel methodology not only allows distinction between positive and negative interdependence, but additionally identifies a yet unexplored form of interaction. This form of interaction covers every type of causality that is neither positive nor negative. Thus, referring to complex forms of interdependencies, we decided to name this broad category as "dark causality", suggesting an unclear way of *X* influencing *Y*.

In order to effectively analyze the spectrum of causal interactions, both transparent (positive, negative) and opaque (dark) in complex systems, we are motivated to establish the general framework on interactions of symbolic dynamics (patterns) (Morse and Hedlund 1938) in reconstructed attractors (Takens 1981). Using basic models, we demonstrate that our method distinguishes between positive, negative, and dark causality. Then, we apply our approach to pairs of financial assets and expose the positive nature of causality between Microsoft and Apple stocks and a competitive interaction between S&P 500 (as proxy of stock market performance) and U.S. government 10-year bond yield. Lastly, we illustrate the prominence of dark causality in the global network of sovereign credit default swaps (CDS).

Ultimately, in Chapter 4, we develop an enhanced version of the pattern causality algorithm for detecting the hidden structure in complex systems beyond financial networks via short-term as well as long-term predictions. Our aim in this chapter was to address the third and final research question regarding the universality of our framework's applications. We effectively demonstrate our method's power to reconstruct the backbone of complex systems by highlighting the most fundamental operations and components. To reveal the multidisciplinary nature of the treatment, we apply our method in three distinct areas of research where we already have an *a priori* knowledge of the crucial operations and components, and thus can reconstruct the most fundamental structure and convincingly evaluate the effectiveness provided by this novel methodology. In this direction, firstly for a desert ecosystem, we capture both the meaningful invasion and subsequent assimilation dynamics of the invader plant species, *Erodium cicutarium,* as well as the effects of drought as charted from precipitation and temperature. Secondly, for a brain activity experiment, we explore and are able to detect an expected (from literature) intensification of activity in the frontal region of the control brain compared to the alcoholic brain, a negative regime in the alcoholic brain between frontal and parietal regions associated with motor functions, as well as higher concentration of activity in the visual cortex of the control brain. Finally, for a set of

banking Credit Default Swaps (CDS), we capture the driving nature of Nordic Banks that is confirmed by the International Monetary Fund, the competitive role of German banks given their balance sheets, as well as the central role of certain banks during the 2007-2008 crisis.

Overall the importance of understanding the hidden structures, not only in the financial markets but also in ecosystems and brain activities, has been progressively coming to a head in public discourse. With the rise and prevalence of big datasets, it is a crucial time for academics to study the interdependency outside conventional analysis. Thus, the novel framework proposed in this thesis does not only allows distinction between positive and negative interdependence, but additionally identifies a yet unexplored form of interaction we coin as “dark causality”.

2 Causality networks of financial assets

The relationship between a variable (the cause), whose past performance influences the future output of another variable (the effect), is known as "causality" (Pearl 2003). Scientists from various disciplines have developed methods to quantify causality in time series, despite the fact that not all of them use this terminology exclusively.

Statisticians, in evolving the notion of correlation, introduced the methods of linear intertemporal cross-correlation (Hawawini 1980) and nonlinear intertemporal cross-correlation (Pijn et al. 1989) in their endeavor to quantify lead–lag relationships in time series. Econometricians, driven by the need to quantify common integrated behavior in time series, developed the methods of linear (Granger 1981) and nonlinear cointegration (Granger and Hallman 1991). Moreover, the known index of Granger causality, in both its linear (Granger 1969) and nonlinear (Hiemstra and Jones 1994) forms, was also established in their field. Last, physicists, mostly from the discipline of information theory, created indexes of mutual information (Granger and Lin 1994) and transfer entropy (Schreiber 2000) that quantify, in a model-free way, the relationships between variables. In this paper, we shall use the above collection of eight causality methods to capture the evolution of average causality in a financial network of assets consisting of various national market indexes, sovereign bonds and oil prices, in the periods before, during and after the global financial crisis.

2.1 Causality networks research motivation

Our motivation to delve into the realm of causality in financial assets is threefold. We want to examine the ability of causality to serve as an early warning signal for systemic financial collapse. We are also interested in exploring whether or not any arbitrage opportunities arise, through, for example, persistent and strong causal relationships between assets. Last, we wish to scrutinize the less explored discipline (compared with correlation) of causalities in financial networks.

Our choice of assets from various national markets rather than from an individual stock market is deliberate. Given that we aim to study the evolution of causality throughout the global financial crisis, we choose both indexes and bonds from countries most representative of the ensuing turmoil. Specifically, we choose market indexes from the United States, China, Brazil, Germany, France, Japan, Hong Kong, Australia and India, as they are countries with large stock market capitalization. For government bonds, we choose the United States, the United

Kingdom, Germany, Japan, Australia, Switzerland, Spain, Italy and Greece, because we consider these governments to have been the most involved in or affected by the crisis. In our analysis, we also include the price of oil: a commodity strong enough to cause a crisis on its own (as in the case of the 1973 oil crisis). As the period we choose for analysis (2000–16) saw extensive use of the Web and social media to enable the international transmission of stock market news and shocks faster than ever before, this necessitates the use of data from many countries. Our study looks at a significant portion of the global financial system during a period of extreme disorder, through the lens of multidisciplinary causality methods.

2.2 Causality networks comparative study

In order to assess whether or not any comparative analysis of the eight causality methods is meaningful, we test the similarity of common links that remain in the financial network (after filtering for the optimal links in terms of the maximum spanning tree) and find that the most similar pair of causality-induced networks is that of linear intertemporal cross-correlation and linear cointegration methods, which score a 48.74% average similarity throughout the time period examined. Thus, we consider it meaningless to try to compare the results of different causality methods, at least through our choice of filtering (maximum spanning tree).

In order to gain portfolio-specific insights, we delve into network analysis, and specifically link persistence and asset centrality. Toward that end we rank the causal links produced by each causality method and find that the most intense and protracted relationship across all causality methods is that of the ten-year US bond causing the prices of the two- to three-year Spanish bond (where price causation is denoted by "→"). The latter relationship can be considered a strong candidate for arbitrage opportunities. Finally, we rank all twenty-five of the financial assets under study in terms of averaged causality issuing to the network by means of out-strength centrality and find that the most causal asset overall seems to be the ten-year US bond. Nevertheless, in general, sovereign bonds ranked better than equities and oil, particularly in the case of linear cointegration and hidden causality, unveiling a hidden regime of bonds. This result could not be better described than in the words of James Carville, President Clinton's political adviser (Arnold 2011), who said: "I used to think that if there was reincarnation, I wanted to come back as the president or the pope.... But now I would like to come back as the bonds market. You can intimidate everybody."

In Section 2.3 we present the formulas and review the literature for each of the eight causality methods and discuss their significance to finance. In Section 2.4, we provide details

regarding our choice of data set. In Section 2.5, we present our research questions and results, and in Section 2.6 we give our concluding remarks.

2.3 Causality Methodologies

In our study, causality between financial assets is quantified as the impact that an asset's past price performance has on another asset's future price performance. Embedded in the nature of causality is some form of predictive potential, i.e., if we know the causality (quantified) between two assets, then, given the price of the cause asset we can, to some extent and probabilistically speaking, forecast the effect asset's price.

2.3.1 Linear Intertemporal cross-correlation

The existence of linear intertemporal cross-correlation (*LICC*) implies that asset prices change in a lead–lag manner and not simultaneously (Atchinson et al. 1987). *LICC* is also known as lead–lag cross-correlation, time-delayed cross-correlation or time-dependent cross-correlation. Hawawini (1980) was the first researcher to implement it in the finance literature, and we present the formula below:

$$LICC_{\Delta t}^{xy} = \frac{\langle R_{\Delta t}^{x}(t) R_{\Delta t}^{y}(t+\tau)\rangle - \langle R_{\Delta t}^{x}(t)\rangle \langle R_{\Delta t}^{y}(t+\tau)\rangle}{\sqrt{\langle [R_{\Delta t}^{x}(t) - \langle R_{\Delta t}^{x}(t)\rangle]^2\rangle \langle [R_{\Delta t}^{y}(t+\tau) - \langle R_{\Delta t}^{y}(t+\tau)\rangle]^2\rangle}}, \tag{2.1}$$

where $R_{\Delta t}^{x}(t) = \log[p(t)] - \log[p(t-\Delta t)]$ is the log return of the price, $p(t)$, of an asset at a certain time t. Δt denotes time interval of the log returns, usually, one-time unit. τ denotes the intertemporal delay among the two assets, and $\langle R_{\Delta t}^{x}(t)\rangle$ denotes the mean of $R_{\Delta t}^{x}(t)$, and x and y denote the two assets.

Remark 2.1. When $\tau = 0$, the LICC coincides with Pearson product moment correlation coefficient. $R_{\Delta t}^{x}(t)$ takes values from -1 to $+1$. When $R_{\Delta t}^{xy}(t) < 0$, this means that asset x has a reverse effect on asset y, ie, if yesterday's return on asset x increases, then today's return on asset y shall decrease, and vice versa. When $R_{\Delta t}^{xy}(t) > 0$, this means that asset x has a same-direction effect on asset y, ie, if yesterday's return on asset x increases then today's return on asset y shall also increase. If $R_{\Delta t}^{xy}(t) = 0$, this means that asset x has no effect on asset y.

The existence of cross-correlations between asset returns implies a deviation from the efficient market hypothesis, and thus provides a probabilistic glimpse at the future asset prices (Atchinson et al. 1987; Lo and MacKinlay 1990; Hawawini 1980).

Kullmann et al. (2002) employed the *LICC* index to analyze a network of equities from the New York Stock Exchange (NYSE). They suggested that the existence of such causal relations is due to the functional interactions between the companies that are represented by the equities in their data set. Mizuno et al. (2004) examined *LICC* in data for foreign exchange rates and pinpointed arbitrage opportunities for the Japanese yen through buying it in one market and selling it in another. Eom et al. (2008) analyzed asset prices from Korea, Japan, Taiwan, Canada and the United States and found statistically significant interactions between them. Wang et al. (2011) analyzed a network of forty-eight stock market indexes and concluded that news regarding stock market crashes travels beyond national boundaries, toppling stock markets around the world in a domino fashion. Huth and Abergel (2014) found important lead–lag relationships in US high-frequency time series. Curme et al. (2015) analyzed time series of 100 NYSE equities and located non-negligible intertemporal cross-correlations, suggesting possible arbitrage opportunities.

The advantage of *LICC* is that it captures the direction of influence between asset returns, unlike the Pearson product moment correlation coefficient, which just captures naive correlations. However, *LICC* has one disadvantage that cannot be ignored: it only takes into account linear causal relationships. It cannot capture nonlinear causal relationships. Therefore, next we present nonlinear intertemporal cross-correlation.

2.3.2 Nonlinear intertemporal cross-correlation

The inability of *LICC* to capture nonlinear intertemporal relations can be overcome by employing the nonlinear intertemporal cross correlation. (*NICC*). This is a statistical measure developed by Pijn et al. (1989) which quantifies both nonlinear as well as linear causality from a time series x to a time series y. Pijn et al. developed *NICC* (also known as "correlation ratio eta") out of need to capture nonlinear time-delayed relationships between neuron signals. We here consider the application of *NICC* to financial time series, since we are interested in capturing the nonlinear intertemporal relationships among assets. According to Pijn et al. (1989), $NICC_{\Delta t}^{xy}$ describes the reduction in uncertainty of $R_{\Delta t}^{y}(t+\tau)$ that can be achieved by forecasting the $R_{\Delta t}^{y}(t+\tau)$ values from those of $R_{\Delta t}^{x}(t)$ via regression as

$NICC_{\Delta t}^{xy} = (totalvariance - unexplained\ variance)/total\ variance$:

$$NICC_{\Delta t}^{xy} = \frac{\sum_{t=1}^{T} R_{\Delta t}^{y}(t+\tau)^2 - \sum_{t=1}^{T} (R_{\Delta t}^{y}(t+\tau) - f(R_{\Delta t}^{x}(t)))^2}{\sum_{t=1}^{T} R_{\Delta t}^{y}(t+\tau)^2}, \quad (2.2)$$

where $f(R_{\Delta t}^{x}(t))$ is the linear piecewise approximation of the nonlinear regression curve.

Remark 2.2 Pijn et al. (1989) commented that, unlike the Pearson product moment correlation coefficient, which is always symmetric, (ie, it is the same for the relationship x, y as for y, x, the *NICC* more often than not is asymmetric $NICC_{\Delta t}^{xy} \neq NICC_{\Delta t}^{yx}$. Interestingly enough, when the relationship f is linear, then *NICC* converges to the same calculation as *LICC*. Note also the fact that, the larger the asymmetry in the values of *NICC* from x to y and vice versa, the more nonlinear the relationship f. *NICC* values move strictly between 0 and 1. $NICC_{\Delta t}^{xy}$ is 0 when y is independent of x and 1 when y is completely determined by x (Pijn et al. 1989).

To the best of our knowledge, *NICC* has never previously been used in the field of finance. It has only been employed in the field of brain signal analysis, in order to determine nonlinear dependencies between neurons (see Pijn et al. 1989, 1990; Lopes da Silva et al. 1989; Wendling et al. 2001). *NICC* is employed to identify nonlinear causal relationships between asset returns. Nonlinearities are important in the finance literature, since many phenomena and relations in finance are nonlinear. Frank and Stengos (1989), after conducting econometric analysis of the returns of gold and silver, found proof that their time series are governed by nonlinear rather than linear mechanisms. Hsieh (1989) analyzed day-by-day variations in major foreign exchange rates through linear correlation and found no significant results. However, after employing econometric generalized autoregressive conditional hetero-scedasticity (GARCH) analysis he identified that nonlinear dependencies saturate the exchange rates under study. Scheinkman and LeBaron (1989) tested dependencies in weekly returns of assets only to realize that no random walk remains in their time series; rather, nonlinear functions better explain those dependencies and also predict future prices from past prices. Abhyankar et al. (1997) examined real time returns of the Standard & Poor's 500 (S&P 500), Deutscher Aktienindex (DAX), Nikkei 225 and Financial Times Stock Exchange 100 (FTSE 100) and found evidence of strong nonlinearities between them.

NICC reveals nonlinearities in the dependencies of asset returns that *LICC* is unable to reveal. This fact renders *NICC* superior to *LICC* in terms of causal relationship detection. However, *NICC*, which takes values from 0 to 1, provides no information about the sign

(positive or negative) of causality. This means that *NICC* cannot tell whether two assets have reverse or same-direction causality, unlike *LICC*, which may not capture nonlinearities but does capture the sign of the causality between asset returns.

2.3.3 Linear cointegration

Lead-lag relationships as examined by *LICC* and *NICC* are one form of causality between assets. Another form of causal relationship is that of assets which move in an integrated way, ie, they evolve dynamically together, and this common evolution can be described by a common function. First, we present the case of assets cointegrated in a linear way. Linear cointegration (*LCo*) is an econometric tool introduced by Granger (1981), and subsequently established by Engle and Granger (1987) and Granger and Weiss (2001). Admittedly, two series can be considered cointegrated when a linear combination of the two is stationary, while neither of the time series is individually stationary (Hakkio and Rush 1989). Following Engle and Granger (1987), we provide the *LCo* method: we must examine whether or not the two series are integrated to the same order. There are various substitution methods to test the integration order of time series: the Dickey–Fuller (Dickey and Fuller 1979), the augmented Dickey– Fuller (ADF) test (Dickey and Fuller 1981), a generalization of the ADF (Phillips and Perron 1988) and the Kwiatkowski–Phillips–Schmidt–Shin test (Kwiatkowski et al. 1992). Given that two series, x_t and y_t, are integrated to the same order, in order to be cointegrated there must exist a function

$$z_t \in I(0): z_t = y_t - \beta x_t. \tag{2.3}$$

For our analysis we shall use the ADF test (Hamilton, 1994). For further technical details, the reader is referred to Engle and Granger (1987). In order to quantify the causal links produced by *LCo* analysis in our dataset, we follow and expand the technique of Yang et al. (2014): assign to every causal relationship of assets the β coefficient from the cointegrating regression and normalize it simply by dividing with the max β coefficient among all asset pairs.

Remark 2.3. Thus, linear normalized cointegration link values can range from -1 to 1. When $LCo_{\Delta t}^{xy} < 0$, this means that asset x is negatively cointegrated with asset y, ie, if yesterday's price of asset x increases, then today's price of asset y will decrease, and vice versa. When, $LCo_{\Delta t}^{xy} > 0$, this means that asset x is positively cointegrated with asset y, ie, if yesterday's

price of asset x increases then today's price of asset y will also increase. If $LCo_{\Delta t}^{xy} = 0$, this means that asset x an asset y are not cointegrated in any way.

In the finance literature, we found that the concept of cointegration (similarly to intertemporal cross-correlation) is linked to the efficient market hypothesis. According to Chowdhury (1991), given that market efficiency ordains that the current asset price dynamically and immediately absorbs and reflects all available information and, given past prices, no further information should increase the predictability of the assets' prices, a cointegration between two financial assets implies inefficiency. Cerchi and Havenner (1988) employed *LCo* analysis in a data set consisting of five randomly chosen industrial stocks. They found that, despite the fact that the individual stock time series could at best be described as random walks, when cointegration enters into play the series appear to have a distinct common trend. Hall et al. (1992) analyzed yields to maturity of US treasury bills and found strong evidence that they move in tandem dynamically through time. Liu et al. (1997) scrutinized the chaotic behavior of the Shanghai Composite Index and Shenzen market indexes. Their analysis shed light on an underlying mechanism between the two indexes, as they seemed to evolve in a cointegrated manner. Alexander (2001) claimed, after conducting robust analysis of commodities, that related commodity types offer some windows of opportunity, given their strong cointegration. Siliverstovs et al. (2005) investigated a data set consisting of natural gas markets in Europe, North America and Japan between the early 1990s and 2004. They found a high level of natural gas market cointegration within Europe, and between the European and Japanese markets as well as within the North American market. Yang et al. (2014) investigated twenty-six stock market indexes and found that their cointegration relationship increased after the Lehman Brothers collapse, while the degree of cointegration gradually decreased from the subprime to the European debt crisis.

LCo is useful enough when we are seeking causality in the sense of assets moving in a linearly integrated manner with the emphasis on a longer temporal horizon than the *LICC*. However, *LCo* is unable to identify nonlinear cointegrating relationships. This is where its nonlinear counterpart enters into play, as described below.

2.3.4 Nonlinear cointegration

Nonlinear cointegration (*NCo*) is an expansion of the well-established linear cointegration (*LCo*) that is capable of capturing nonlinear integrated dependencies between one asset and another. *NCo* was introduced by Granger and Hallman (1991), and further developed by Balke

and Fomby (1997), Escribano and Mira (2002) and Escanciano and Escribano (2011). According to Escanciano and Escribano (2011), two "extended memory" series y_t and x_t are nonlinearly cointegrated if there exists a function f such that:

$$z_t = f(y_t, x_t) \text{ , is short in memory.} \tag{2.4}$$

Crashes in extended memory time series have an enduring and intense impact; while in short memory series crashes are absorbed and vanish quickly (Escanciano and Escribano, 2011). Memory in time series, and its characterization as short or extended, can be measured via various means. In our analysis, we use the conditional mean persistence method from the Escanciano and Escribano (2011). A time series x_t is considered to be of "short memory in mean" if, for all t and $h > 0$, $M(t,h) = E(y_{t+h} \mid I_t), h > 0$, tends to a constant μ as h becomes large (for more details see Escanciano and Escribano, 2011). Given that we found no method of quantifying the nonlinear cointegration among two variables in the literature, we devise our own method which is described below.

We assign the weighted average of the coefficients in function f from (2.4) to be the weight of a nonlinear cointegration from asset x to asset y. We allow our algorithm to search for candidate functions f up to tenth-degree polynomials, thus the higher the term's power the greater the weight assigned to it. For each cointegrating relationship, we divide the sum of these ten coefficients by the maximum of the coefficient's averages between all asset pairs, in order to claim a normalized quantity for NCo. Thus, nonlinear cointegration link values (as normalized by us) can range from -1 to 1. If $NCo^{xy} < 0$, this means that asset x is negatively cointegrated with asset y. When $NCo^{xy} > 0$, this means that asset x is positively cointegrated with asset y. If $NCo_t^{xy} = 0$, asset x and asset y are not cointegrated in any way.

Li (2002) analyzed the stock market indexes of Australia, Japan, New Zealand, the United Kingdom and the United States in terms of both *LCo* and *NCo*. His results indicated that nonlinear cointegration relationships between those indexes are much stronger and persistent than the linear ones. Ma and Kanas (2004) found further strong empirical evidence to support the intrinsic bubble model of stock prices developed by Froot and Obstfeld (1991). Athanasenas et al. (2014) conducted analysis of the time series of revenues and expenditures of the Greek government. Their results support the fact that negative rates of expenditure severely affect revenues. Apergis and Payne (2014) analyzed asset returns from the stock

markets of the Group of Seven (G7). They found long-lasting nonlinear dependencies in a significant portion of their data set.

2.3.5 Linear Granger causality

Granger causality is a statistical concept of causality that is based on regression. It has been widely used in the financial econometrics literature to detect causal relationships among assets and other economic variables. According to Granger causality, if a time series x_t "Granger-causes" (or "G-causes") a time series y_t, then past values of x_t should contain predictive information that serves to forecast y_t better than the information contained in past values of y_t alone. The so-called predictive information is modelled through regression, (linear regression for linear Granger causality (*LGC*), and nonlinear regression for nonlinear Granger causality (*NGC*) in the next section) Following Granger (1969), given x_t and y_t are stationary, we can consider a linear autoregressive (AR) model of time series y_t:

$$y_t = \sum_{i=1}^{N} a_{11,i} y_{t-i} + \sum_{i=1}^{N} a_{12,i} x_{t-i} + E_t(y), \tag{2.5}$$

where N is the number of past observations included in the AR model, $a_{11,i}$ and $a_{12,i}$ are the coefficients of the model, $E_t(y)$ is the residual, also known as prediction errors, for y_t. We can say that $x\ G-causes\ y$ if and only if the coefficients $a_{12,i}$ are significantly different from zero. To test the underlying significance, we employ the F-test with the null hypothesis that $a_{12,i} = 0$. In the literature we found no method of quantifying the weight of a G-causal link from an asset x to an asset y.

So given that asset x G-causes y, we decide to assign as weight of the link the value: $LGC_{xy} = 1 - \mathrm{p} - value$ of the F-test. Values range from 0 to 1.

Bradshaw and Orden (1990) uncovered an important *LGC* between the exchange rate and export sales, while the evidence on causality of the exchange rate on prices is unclear. Rahman and Mustafa (1997) analyzed US equities and bonds. Their results attested that the causality from bonds to equities is much more robust than from equities to bonds. Abhyankar (1998) investigated causal relationships between futures contracts and cash markets. According to his results, futures contracts hold predictive information for the future states of the cash market. Dutta (2001) found that the causality from levels of telecommunications infrastructure to economic activity is stronger than that in the opposite direction. Foresti (2006) scrutinized the

possibility of causal relationships between economic growth and stock market returns, concluding that the stock market drives economic growth. Wang et al. (2007) tested for possible linkages between the euro and US, Japanese and German interest rates. Their results indicated that Japanese interest rates exert intense causality overall, and that the German interest rates have a bidirectional causal relationship with a variety of euro currency rates. Zhang and Wei (2010) found that crude oil prices have a statistically significant causal relationship to the prices of gold, but that the opposite was not supported. Billio et al. (2012) analyzed time series data of hedge funds, banks, broker-dealers and insurance companies and found that banks exert the most causality on all the other time series they analyzed. Výrost et al. (2015) uncovered an underlying mechanism of a preferential attachment between stock markets, ie, the probability of spillover effects between any given two markets increases with their degree of connectedness to other markets. Fiedor (2015) analyzed the relationships between companies listed on the S&P 100 and found that causal relationships are more prevalent than lagged synchronization relationships.

One drawback of *LGC* is that it does not provide any information regarding whether the assets under study have positive or negative causality. Another is its inability to capture nonlinearities. The latter drawback is avoided by its nonlinear counterpart, which is presented below.

2.3.6 Nonlinear Granger causality

Nonlinear Granger causality (*NGC*) is capable of mining the nonlinear predictive information that a time series can hold about another time series, a feat that *LGC* fails to accomplish. *NGC* was introduced by Hiemstra and Jones (1994), and further established by Péguin-Feissolle et al. (2013). Following the definition of Péguin-Feissolle et al., we let y_t and x_t be two stationary and ergodic time series. In order to test the existence of a causal relationship between two series, we define:

$$y_t = f_y(y_{t-1}, y_{t-p_1}, x_{t-1}, x_{t-q_1}; 1) + e_{1,t}. \tag{2.6}$$

This equation includes all combinations between past values of y and x. We can say that x nonlinearly G-causes y if and only if the coefficients on the terms of x's past values are significantly different from zero. To test the underlying significance, we can use the Wald F-

test (for more technical details regarding the methodology of NGC see Péguin-Feissolle et al., 2013). In the literature we found no method of quantifying the weight of a nonlinear G-causal link from an asset x to an asset y. So given that asset x nonlinearly Granger causes y, we decided to assign as weight of the link the value: $NGC_{xy} = 1 - P\ value$ of the Wald F-test. Values denoting the intensity of the causality range from 0 to 1.

Hiemstra and Jones (1994), having used the *NGC* between the equity returns of DJIA stocks and the volue rate of the NYSE, found statistically significant causality in both directions. Qiao and Lam (2011) sought causalities in East Asia stock markets. Despite the fact that *LGC* tests failed to detect statistically significant dependencies, its nonlinear relative, *NGC*, captured many causalities. Benhmad (2012) similarly analyzed oil prices and the US dollar exchange rate and found evidence of oil influencing the US dollar rate slightly more than the other way around. Zhou et al. (2014) examined a data set of Chinese stock futures and spot markets. They claim to have found robust results in favor of futures influencing spot markets. Chu et al. (2016) researched equity returns and investor sentiment in China. Surprisingly enough, they found that both types of time series influence each other in a nonlinear way.

In the last two sections on causality tools below, we deviate from the disciplines of statistics and econometrics and recall methodologies from information theory.

2.3.7 Shadow causality

The causality methods described above have several important weaknesses: they depend on requirements of stationarity (*LICC, NICC, LCo, NCo, LGC, NGC*), are unable to capture nonlinearities (*LICC, LCo, LGC*) or they cannot distinguish between positive (homogeneous) and negative (heterogeneous) causality (*NICC, LGC, NGC*). This is where two tools from information theory come into play: shadow causality (*SC*), which is based on mutual information; and hidden causality (*HC*), which is based on transfer entropy. These information-theoretic methods are nonparametric, have no requirements of stationary time series and capture both linear and nonlinear causality. With a minor modification, we were also able to make them distinguish between positive and negative causality. Following the "shadow correlation" of Granger and Lin (1994), and inspired by Schreiber (2000), who suggested that a lead/lag can be used in mutual information to include directionality in the calculations, we give the shadow causality formula as follows:

$$SC_{x_{t-\Delta t},y_t} = sgn\sqrt{1 - exp^{(-2I(x_{t-\Delta t},y_t))}}, \tag{2.7}$$

Where

$$I(x_{t-\Delta t}, y_t) = \int \int p_{x,y}(x_{t-\Delta t}, y_t) \log \frac{p_{x,y}(x_{t-\Delta t}, y_t)}{p_x(x_{t-\Delta t})\, p_y(y_t)} dx\ dy$$

is the mutual information (MI). $p(x, y)$ denotes the joint probability distribution function of time series x_t and y_t. $p(x)$ and $p(y)$ denote the marginal probability distribution functions of x_t and y_t, respectively. The function $SC_{x_{t-\Delta t},y_t}$ captures the overall linear and nonlinear causality from x to y. If causality is homogeneous, then $sgn = +1$ (homogeneous causality takes place when an increase in x causes increase in y more often than decrease in y while at the same time decrease in x causes decrease in y more often than increase in y). If causality is heterogeneous, then $sgn = -1$ (heterogeneous causality takes place when increase in x causes decrease in y more often than increase in y while at the same time decrease in x causes increase in y more often than decrease in y). *SC* values denoting the intensity and type of the causality range from –1 to 1 (negative values denote a heterogeneous relationship and positive values denote a homogeneous relationship).

Dionisio et al. (2007) created a bundle of economic and financial indicators for Portugal and tested for possible dependencies between them via *MI* analysis. Their analysis showed that there are strong causalities from dividend yield and earnings price ratio time series to the monthly excess returns of investors. Maasoumi and Wang (2008) tested for dependencies in economic growth time series between various municipalities in China. They found significant formations of groups between municipalities, manifesting before and after reformation periods. Menezes et al. (2012) analyzed equity time series, representative of the G7 countries. Comparing their results with other methods, such as *LGC*, they claim that *MI* provided more information regarding the underlying causalities in the stocks of their data set. Fiedor (2014a) investigated nonlinear relationships between companies listed in the NYSE and found that the mutual information rate produces different results than simple correlation. In the next section, we present our last causality tool, concluding our methodology section.

2.3.8 Hidden causality

Mutual information, which is the basis of shadow causality, needs a time lag to account for directionality and thus causality. Therefore, we can argue that it is not a natural tool of causality interference (much like *LICC* and *NICC*) but rather a manufactured one. Transfer entropy (*TE*) on the other hand, which is the basis of *HC*, is a natural tool of causality inference. *TE* is one of the youngest members of the causality family, as it was only recently introduced by Schreiber (2000). It exploits past values of time series x_t and y_t to test their predictive power for the future value of y_{t+1}. In a similar way to SC, we introduce at this point the HC method which instead of MI uses the stricter TE (Schreiber, 2000). Thus, HC is the normalized version of TE, and for its normalization technique we used the method by Sandoval (2014). Our contribution lies only in the part of mining the sign of positive or negative in a similar manner as we did in SC. The HC formula is given by:

$$HC_{x_t \to y_t} = sgn \ \frac{\sum p(y_{t+1}, y_t, x_t) \log \frac{p(y_{t+1}\, y_t x_t)}{(p(y_{t+1} y_t)}}{\sum p(y_t, x_t) \log_2 \frac{p(y_t, x_t)}{p(x_t)}}, \tag{2.8}$$

Where

$$\sum p(y_{t+1}, y_t, x_t) \log \frac{p(y_{t+1}\, y_t x_t)}{p(y_{t+1}\, y_t)} = TE_{x_t \to y_t},$$

which is the TE of x to y and

$$\sum p(y_{t+1}, y_t) \log_2 \frac{p(y_{t+1}, y_t)}{p(y_t)} = H_{y^F | y^P} \ ,$$

which is the conditional entropy of the future of y on its past. For more technical details; see the papers of Schreiber (2000) and Junior (2013). If causality is homogeneous, then $sgn = +1$ (homogeneous causality takes place when an increase in x causes increase in y more often than decrease in y while a decrease in x causes decrease in y more often than increase in y). If causality is heterogeneous, then $sgn = -1$ (heterogeneous causality takes place when increase in x causes decrease in y more often than increase in y while at the same time decrease in x causes increase in y more often than decrease in y). HC value denoting the intensity and type of causality range from -1 to 1 (negative values denote a heterogeneous relationship and positive values denote a homogeneous relationship).

Baek et al. (2005) analyzed via *TE* a data set consisting of equities from various industrial sectors and found that energy – related equities such as oil, gas and electricity saturate the whole market. Kwon and Yang (2008) sought causal relationships in international stock market indexes and discovered that S&P 500 has the highest number of causal relationships with all the other indexes. Kim et al. (2013) examined stock market indexes for most of the Group of Twenty. Their results stand in favor of the theory that western countries exert stronger causality on eastern countries than vice versa. Sandoval (2014) scrutinized the companies in the S&P 100. His results indicate that *TE* produces a network that creates much more realistic (in terms of industrial affinity) clusters than *LICC*. Sandoval et al. (2015) analyzed the pairs of eighty – three stock market indexes in various countries and found that *TE* is an effective way to quantify the information flow between indexes. Yook et al. (2016) studied a financial network and found that the modular structure from *LICC* does not correctly reflect the known industrial classification and its hierarchy, unlike the transfer entropy method, which fits the market segmentation much better.

Table 2.1 briefly summarizes all the causality methods described in Section 2.3, and some of their basic properties.

Table 2.1: Causality methods: data.

Causality method	**Type of causality identification**	**Value range**	**Needs time series to be stationary**
LICC	Linear	[-1, 1]	Yes
NICC	Linear and nonlinear	[0, 1]	Yes
LCo	Linear	[-1, 1]	Yes
NCo	Nonlinear	[-1, 1]	Yes
LGC	Linear	[0, 1]	Yes
NGC	Nonlinear	[0, 1]	Yes
SC	Linear and nonlinear	[-1, 1]	Yes
HC	Linear and nonlinear	[-1, 1]	Yes

LICC, linear intertemporal cross-correlation. *NICC*, nonlinear intertemporal cross-correlation. *LGC*, linear Granger causality. *NGC*, nonlinear Granger causality

2.4 Data and filtering

For our analysis, we use weekly data for stock market indexes, sovereign bonds and oil from Thomson Reuters DataStream for the period from January 4, 2000 to February 12, 2016. By using weekly data, we negate the time zone effects due to the different operating hours of the stock exchanges in different countries. The idea is to have a broad and global selection of financial assets, and to understand their interactions over time by means of causality analysis. Thus, our data set consists of the following (see Table 2.2).

- Ten stock market indexes: Shanghai Composite Index, Bovespa, DJIA, S&P 500, DAX 30, Hang Seng, CAC 40, Nikkei 225, ASX 200 and the Bombay Stock Exchange (BSE 30 Sensex).
- Fourteen bonds: two-year US bond; ten-year US bond; ten-year UK bond; two-year German bond; ten-year German bond; two-year Japanese bond; ten-year Japanese bond; two-year Australian bond; ten-year Australian bond; ten-year Swiss bond; two- to three-year Spanish bond; ten-year Greek bond; three-year Italian bond and ten-year Italian bond.
- Oil.

We use a rolling window of two years to construct the evolutionary financial network, which evolves week by week from January 4, 2000 until February 12, 2016 (a total of 738 weeks of network evolution) for each of the eight causality tools presented in Section 2.3. The time lag used for our analysis is one week, and the statistical significance of keeping a causal link is 95%. In order to negate nonstationarity we take the log returns of the time series and test with ADF for nonstationarity. No time series in our data set is found to be nonstationary when log returns apply. After constructing 25 x 25 matrixes for each of the 738 weeks and each of the eight causality methods, we apply the filtering method of the maximum spanning tree to each matrix (Hu 1961). Thus, we are able to apply the strongest known filtering and keep only the most powerful causal relations. Nevertheless, it is quite common to see other filtering methods applied to financial networks, such as the minimum spanning tree (MST), which has been thoroughly employed in the works of Mantegna (1999 a, b), who was the first to introduce it in finance. The MST was also effectively used by Di Matteo et al. (2010) and Aste and Di Matteo (2006), who also added their own flavor to filtering. Other filtering methods are the random matrix theory used by Iori et al. (2007); Bonferroni statistical filtering, which was well presented by Iori et al. (2015); and the planar maximally filtered graph, which has been applied

in the works of Di Matteo et al. (2010), Kenett et al. (2010), Birch et al. (2015) and Musmeci et al. (2015, 2016 a, b)

Table 2.2: Data set details and asset numbering for Figures 2.9–2.16. (Equities)

#	Asset	Details
1	Shanghai Composite Index	A stock market index of all stocks (A shares and B shares) traded at the Shanghai Composite Index Stock Exchange
2	Bovespa	An index of about fifty stocks traded on the São Paulo Stock, Mercantile and Futures Exchange
3	DJIA	A stock market index, one of several created by Wall Street Journal editor and DJIA & Company cofounder Charles Dow
4	S&P 500	A US stock market index based on the market capitalizations of 500 large companies having common stock listed on the NYSE or Nasdaq
5	DAX 30	A blue chip stock market index consisting of the thirty major German companies trading on the Frankfurt Stock Exchange
6	Hang Seng	A free-float-adjusted market capitalization-weighted stock market index in Hong Kong, used to record and monitor daily changes in the largest companies on the Hong Kong stock market
7	CAC 40	A capitalization-weighted measure of the forty most significant values of the 100 highest market caps on the Euronext Paris
8	Nikkei 225	A price-weighted index of the Osaka Stock Exchange, with components reviewed once a year; the Nikkei is the most widely quoted average of Japanese equities
9	ASX 200	A market-capitalization-weighted and float-adjusted stock market index of Australian stocks listed on the Australian Securities Exchange from S&P
10	BSE 30 Sensex	A free-float market-weighted stock market index of thirty well-established and financially sound companies listed on the Bombay Stock Exchange

TABLE 2.2. (cont.): Data set details and asset numbering for Figures 2.9–2.16. (Bonds and Oil)

#	Asset	Details
11	2Y US bond	A two-years-to-maturity sovereign US bond
12	10Y US bond	A ten-years-to-maturity sovereign US bond
13	10Y UK bond	A ten-years-to-maturity sovereign UK bond
14	2Y German bond	A two-years-to-maturity sovereign German bond
15	10Y German bond	A ten-years-to-maturity sovereign German bond
16	2Y Japanese bond	A two-years-to-maturity sovereign Japanese bond
17	10Y Japanese bond	A ten-years-to-maturity sovereign Japanese bond
18	2Y Australian bond	A two-years-to-maturity sovereign Australian bond
19	10Y Australian bond	A ten-years-to-maturity sovereign Australian bond
20	10Y Swiss bond	A ten-years-to-maturity sovereign Swiss bond
21	2to3Y Spanish bond	A two- to three-years-to-maturity sovereign Spanish bond
22	10Y Greek bond	A ten-years-to-maturity sovereign Greek bond
23	3Y Italian bond	A three-years-to-maturity sovereign Italian bond
24	10Y Italian bond	A ten-years-to-maturity sovereign Italian bond
25	Oil	Crude oil as traded in the New York Mercantile Exchange

2.5 Causality network analytics

2.5.1 Gauging causality in turbulent times

As we are motivated to examine the early warning capacity of causalities for financial turbulence, we probe how the financial network changes over time. Was the global financial crisis somehow imprinted on the average causality of the market before, during or after the event? In order to seek answers, we examined the evolution of the average causality in the network, week by week, in a rolling window of one-year length.

During the bursting of the dot-com bubble in the early 2000s, the *LICC* in Figure 2.1 is on average 33% and displays a slightly upward trend as it peaks on October 2002. Before the crisis, it moves at an average of 27.42%; more specifically, it undergoes a marked downward slide throughout 2003, and then from 2004 to July 2007 it fluctuates, with a slight upward trend (Table 2.3). The global financial crisis period is characterized by a dramatic drop in *LICC,* with an average value of 24.29%. Then, in the summer of 2009, we observe a confident rising move as the echoes of the crisis dwindle.[1] The post-crisis period is characterized by a generic and oscillating drop in *LICC,* which is on average 18.95% (Table 2.4). *LICC* enters the recent financial crash of China in an upward trend; however, its levels are already as low as 15.08% (Table 2.5).

As we observe the evolutionary behavior of the *NICC* (see Figure 2.2) during the stock market downturn of the early 2000s, it is on average 29.18% and remains flat throughout that decade. During the pre-crisis period it moves at an average of 28.33%; more specifically the *NICC* moves slightly downward from 2004 to 2006, until it suddenly drops further, half a year before the emergence of the global financial crisis (Table 2.3).

The global financial crisis period is characterized by a dramatic increase in *NICC,* with an average value of 40.91%, but then in summer 2009 we observe a plunge until the crisis dies out. The postcrisis period is characterized by a marked increase in *NICC,* which is on average 49.09% (Table 2.4). The *NICC* enters the stock market plunge of China in an upward trend; however, its levels are already as low as 19.38% (Table 2.5).

Regarding *LCo,* as seen in Figure 2.3, during the dot-com bubble burst it is on average 20.34% and displays an upward trend throughout the downturn (October 2002). During the precrisis period *LCo* displays extreme fluctuations around an average of 26.66% and just before the crisis its trend transforms to a rising one (Table 2.3).

The global financial crisis period is characterized by a smooth drop in *LCo,* with an average value of 30.06%, but then in summer 2009 it stops falling and moves on a marked support level around 25%. The postcrisis period is characterized by a continued move on the same resistance level as before until the end of 2011. Then *LCo* spikes around summer 2012, at 40%, and begins a marked downward trend at an average of 24.04% (Table 2.4). *LCo* meets the financial crisis

[1] Here and later, "summer" is the period from June to August in the northern hemisphere

caused by China in a change of trend from negative levels to positive ones; however, its levels are already as low as - 0.53% (Table 2.5).

As far as *NCo* is concerned (Figure 2.4), during the stock market crash of the early 2000s it is on the negative side, averaging - 6.03%, and displays a slightly upward trend as it peaks during the last gasp of the downturn (October 2002), nearing the zero level. During the precrisis period *NCo* moves at an average of 0.44%; more specifically, it displays an abrupt spike in the first half of 2003, hitting a ceiling of 20%, and then from 2004 to July 2007 it fluctuates with a marked downward trend, joining the negative side again as early as 2006. It bottoms out just before the breakout of the crisis at the lowest level ever, around –15% (Table 2.3). *NCo* enters the Chinese market crash after a prolonged trail with an attractor around zero, being sometimes positive and sometimes negative but always averaging 0.18% (Table 2.5).

Next, we focus on *LGC* (Figure 2.5), and note that during the downturn of the early 2000s it moves around a support level of 77.52%. During the precrisis period it fluctuates a little higher than before at an average of 83.81%. After 2004, however, *LGC* is characterized by a smooth upward trend, until the last quarter of 2006, when it starts rolling somewhat downward (Table 2.3). The outbreak of the global financial crisis is characterized by a faintly diminishing *LGC* with an average value of 74.8%, but then, at the end of 2008, we observe a marked drop below 70%, reaching a new support level at the end of the global crisis. The postcrisis period is characterized by a generic and fluctuating increase in *LGC*, which is on average 64.62%, until summer 2012. However, after the third quarter of 2012 the trend changes to a diminishing one, hitting as low as 38% (Table 2.4). *LGC* enters the Chinese downturn in an upward trend; however, its levels are already as low as 53.62% (Table 2.5).

Concerning *NGC* (Figure 2.6), emerging from the dot-com burst it has a faintly upward trend, averaging 68.81%, becoming steeper as the downturn dies out in October 2002. During the precrisis period *NGC* plateaus at an average of 77.96%; this lasts until the end of 2004, when it suddenly drops to a support level of 60%. Then, in summer 2005 it bounces back up, continuing its rising trend until summer 2006, just one year before the outbreak of the crisis, when it reaches a resistance level of 88% and begins falling again (Table 2.3). On the eve of the global financial crisis *NGC* falls again, to the same support level as before (60%), while at the end of 2008 it bounces up to between 70% and 80%, moving at an average value of 73.14% until the end of the crisis. The postcrisis period is characterized by a historically unique drop in *NGC*, reaching a deep support level of just 20% and staying there from summer 2012 until

summer 2013, when it starts to rise again in a volatile manner (Table 2.4). The late Chinese shock finds *NGC* still on the rise, at an average of 59.41% (Table 2.5).

Furthermore, the post-dot-com bubble burst is characterized by an *SC* (Figure 2.7), on average 18.77%, while its trend is downward. During the precrisis period *SC* enters a marked rising trend, averaging 32.16% even after the birth of the global crisis. Throughout the precrisis period *SC* more than doubles, from 20% in 2003 to 50% in summer 2007 (Table 2.3). The global financial crisis period is characterized by an extreme plunge in *SC,* from the resistance level of 50% to a support level of 17%, with an average value of 29.97%. In the twilight of the crisis (2009) *SC* plateaus at 20%; this continues into the postcrisis period. The postcrisis period is characterized by a marked increase in *SC* to a new support level of 30% between 2011 and 2012, and then in 2013 this support level becomes a resistance level, forcing the *SC* to stay beneath 30%, with an average of 28.65% (Table 2.4). The *SC* enters the financial crash of China in a faintly downward trend, and then plateaus at 28.31% (Table 2.5).

Ultimately, during the stock market downturn of the early 2000s, the *HC* (Figure 2.8) is on average 5.48% and displays a slightly upward trend as it peaks on the exhaustion of the downturn (October 2002). During the precrisis period it moves at an average of 14.19%; more specifically, the HC is governed by a steep increase from 2003 through summer 2004, when it hits the resistance level of 25% and then enters a prolonged fluctuation with a downward trend to the support level of 10% in summer 2006, almost one year before the global crisis (Table 2.3). The global financial crisis finds *HC* on the rise, with an average value of 11.90%, but then in summer 2008 we witness a marked diminishing trend as the global crisis enters its mature phase. Then we can see that the *HC,* having fallen to the negative side at around 5%, rallies upward by 15% at the end of 2012, and fluctuates above the support level of 5% until 2014, when it again makes a steep decline to subzero levels (Table 2.4). The *HC* enters the financial crisis caused by China in a slightly upward trend; however, its levels are already as low as 5.35% (Table 2.5).

Figure 2.1: Average of *LICC* for all assets week by week, with a rolling window of two years. Shading denotes periods of financial turmoil: 2001-2, post-dot-com bubble burst; 2007-9, global financial crisis; 2015, Chinese stock market crash.

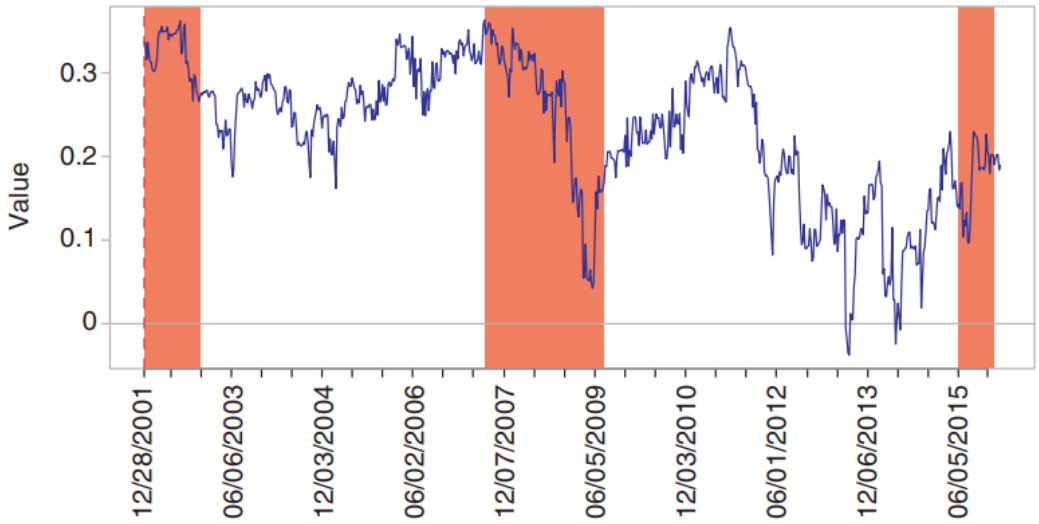


Figure 2.2: Average of *NICC* for all assets week by week, with a rolling window of two years. Shading denotes periods of financial turmoil: 2001–2, post-dot-com bubble burst; 2007–9, global financial crisis; 2015, Chinese stock market crash.

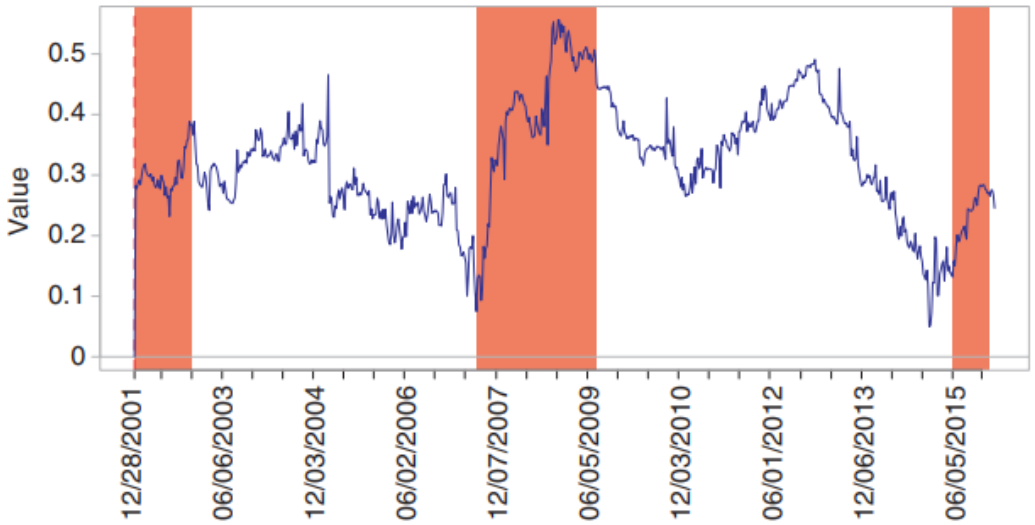


Figure 2.3: Average of *LCo* for all assets week by week, with a rolling window of two years. Shading denotes periods of financial turmoil: 2001–2, post-dot-com bubble burst; 2007–9, global financial crisis; 2015, Chinese stock market crash.

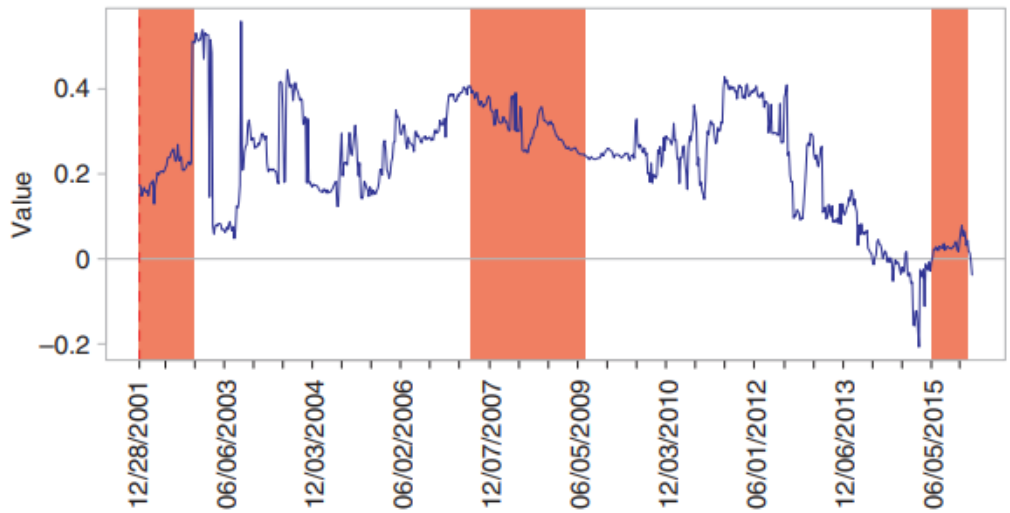


Figure 2.4: Average of *NCo* for all assets week by week, with a rolling window of two years. Shading denotes periods of financial turmoil: 2001-2, post-dot-com bubble burst; 2007-9, global financial crisis; 2015, Chinese stock market crash.

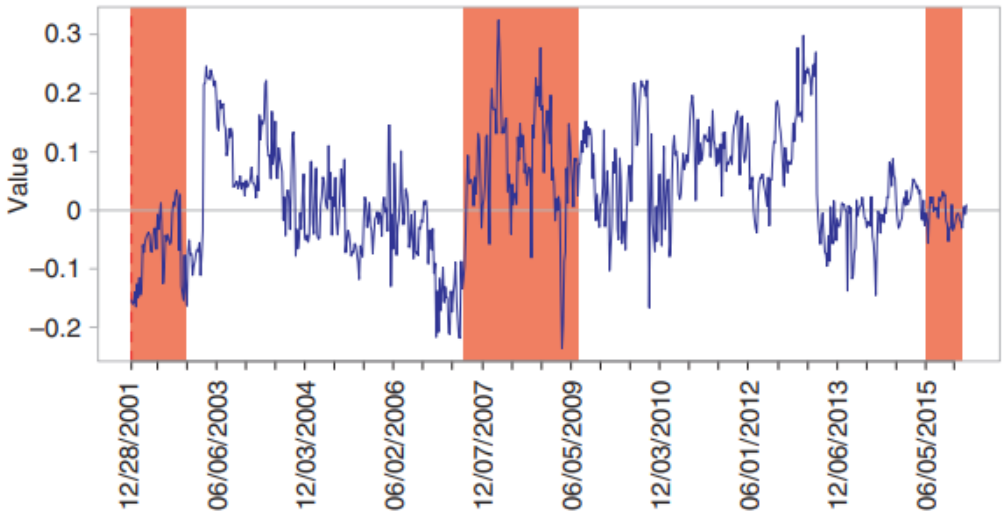


Figure 2.5: Average of *LGC* for all assets week by week, with a rolling window of two Shading denotes periods of financial turmoil: 2001-2, post-dot-com bubble burst; 2007-9, global financial crisis; 2015, Chinese stock market crash.

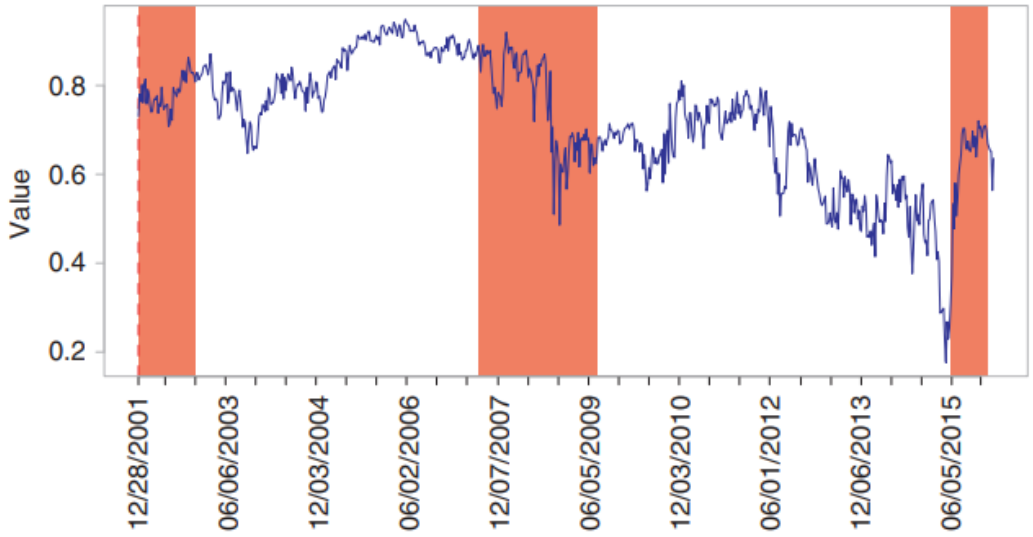


Figure 2.6: Average of *NGC* for all assets week by week, with a rolling window of two years. Shading denotes periods of financial turmoil: 2001-2, post-dot-com bubble burst; 2007-9, global financial crisis; 2015, Chinese stock market crash.

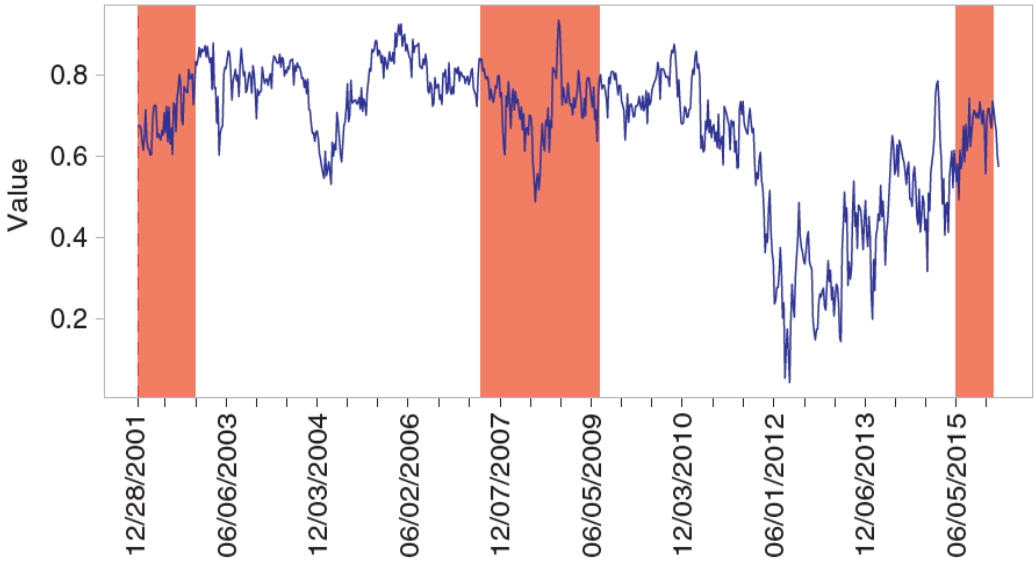


Figure 2.7: Average of SC for all assets week by week, with a rolling window of two years. Shading denotes periods of financial turmoil: 2001-2, post-dot-com bubble burst; 2007-9, global financial crisis; 2015, Chinese stock market crash.

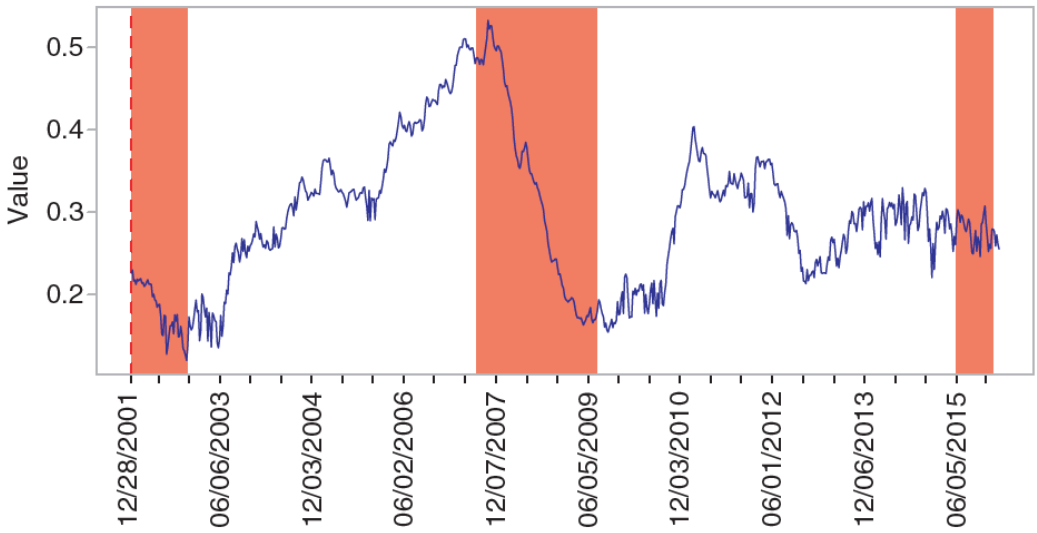


Figure 2.8: Average of HC for all assets week by week, with a rolling window of two years. Shading denotes periods of financial turmoil: 2001-2, post-dot-com bubble burst; 2007-9, global financial crisis; 2015, Chinese stock market crash.

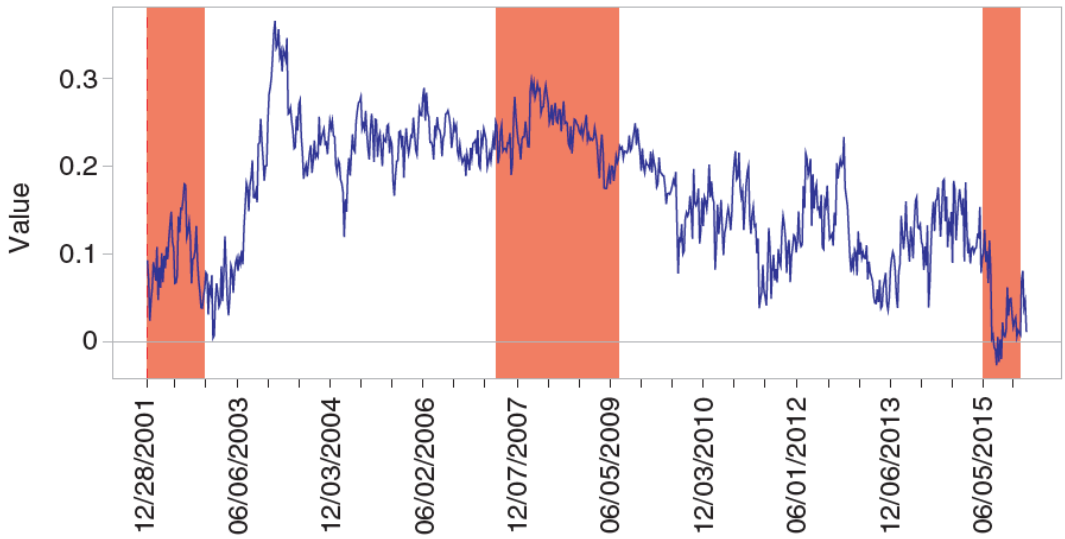


Table 2.3: Causalities: general statistics from the dot-com bubble burst until before the global financial crisis.

Causality methods	**Stock market downturn of post-dot-com bubble burst: Mar 2000 to Oct 2002**				
	Min	**Avg**	**Max**	**SD**	**Trend**
LICC	0.2666	0.3300	0.3620	0.232	→
NICC	0.2320	0.2918	0.3469	0.0188	→
LCo	0.1294	0.2034	0.2692	0.0339	↑
NCo	-0.1650	-0.0603	0.0348	0.0564	↑
LGC	0.7076	0.7755	0.8631	0.0333	→
NGC	0.6026	0.6881	0.8138	0.0562	↑
SC	0.1278	0.1877	0.2297	0.0281	↓
HC	0.0042	0.0548	0.0903	0.0179	↑
Causality methods	**Precrisis period: Nov 2002 to Jul 2007**				
	Min	**Avg**	**Max**	**SD**	**Trend**
LICC	0.1618	0.2742	0.3584	0.0372	↓ → ↑
NICC	0.1008	0.2833	0.4655	0.0612	→ ↑
LCo	0.0478	0.2666	0.5589	0.1164	↓ ↑ ↑ → ↓
NCo	-0.2186	0.0044	0.2452	0.1008	↓
LGC	0.6472	0.8381	0.9485	0.0684	→ ↓ →
NGC	0.5313	0.7796	0.9246	0.0785	→ ↑ ↓ →
SC	0.1203	0.3216	0.5103	0.0965	↑
HC	0.0225	0.1419	0.3409	0.0616	↑ ↓ ↑ ↓

LICC, linear intertemporal cross-correlation. *NICC*, nonlinear intertemporal cross-correlation. *LCo*, linear cointegration. *NCo*, nonlinear cointegration. *LGC*, linear Granger causality. *NGC*, nonlinear Granger causality. *SC*, shadow causality. *HC*, hidden causality. Minimum, average, maximum and standard deviation (SD) are calculated in terms of the average causality of all financial assets throughout the time period declared. Trend symbols: ↑ upward trend; ↓ downward trend; → flat. When two or more trend symbols are written in a row they symbolize consecutive trend changes.

Table 2.4: Causalities: general statistics during and after the global financial crisis.

Causality methods	**Global financial crisis: post-dot-com bubble burst: Aug 2007 to Dec 2009**				
	Min	**Avg**	**Max**	**SD**	**Trend**
LICC	0.0425	0.2429	0.3632	0.0844	↓ ↑
NICC	0.0748	0.4091	0.5563	0.1053	↑ ↓
LCo	0.2336	0.3006	0.4065	0.0498	↓ →
NCo	-0.2359	0.0771	0.3241	0.0900	↑ ↓ ↑ ↓ ↑ ↓
LGC	0.4861	0.7480	0.9192	0.0989	→ ↓ →
NGC	0.4887	0.7314	0.9344	0.0752	↓ ↑ →
SC	0.1545	0.2997	0.5327	0.1269	↓
HC	0.0310	0.1190	0.1996	0.0353	↓
Causality methods	**Postcrisis period: Jan 2010 to May 2015**				
	Min	**Avg**	**Max**	**SD**	**Trend**
LICC	0.0377	0.1895	0.3540	0.0903	↑ ↓ ↑
NICC	0.2379	0.3610	0.4909	0.0609	↑ ↓
LCo	-0.0134	0.2404	0.4281	0.1080	→ ↑ ↓
NCo	-0.1667	0.0679	0.2979	0.0890	↑ → ↓ ↑ ↓ → ↑ →
LGC	0.4148	0.6462	0.8098	0.0952	→ ↓
NGC	0.0441	0.5327	0.8754	0.2073	↓ ↑
SC	0.1717	0.2865	0.4037	0.0559	↑ → ↓ →
HC	-0.0887	0.0534	0.1625	0.0649	↓↑ ↓ ↑

For notation see Table 2.3.

Table 2.5: Causalities: general statistics during the Chinese stock market crash.

Causality methods	**Global financial crisis: post-dot-com bubble burst: Aug 2007 to Dec 2009**				
	Min	**Avg**	**Max**	**SD**	**Trend**
LICC	0.0187	0.1508	0.2298	0.0495	↑
NICC	0.0499	0.1938	0.2839	0.0564	↑
LCo	-0.2074	-0.0053	0.0782	0.0525	→
NCo	-0.1452	0.0018	0.0881	0.0357	→
LGC	0.1763	0.5362	0.7201	0.1349	↑
NGC	0.3172	0.5941	0.7849	0.0998	↑
SC	0.2208	0.2831	0.3296	0.0220	→
HC	0.0535	0.0685	0.1730	0.0642	↓→ ↑

For notation see Table 2.3

Table 2.6: Averaged similarity regarding the link structure after the MST filtering throughout the time period.

Causality methods	***LICC***	***NICC***	***LCo***	***NCo***	***LGC***	***NGC***	***SC***	***HC***
LICC	100	3.12	1.07	5.54	48.74	29.69	3.41	15.77
NICC	3.12	100	1.67	8.33	7.67	6.65	3.84	3.28
LCo	1.07	1.67	100	1.98	1.94	2.27	4.27	4.78
NCo	5.54	8.33	1.98	100	4.54	4.04	4.47	2.71
LGC	48.74	7.67	1.94	4.54	100	39.44	2.24	15.637
NGC	29.69	6.65	2.27	4.04	39.44	100	2.56	13.01
SC	3.41	3.84	4.27	4.47	2.24	2.56	100	4.36
HC	15.77	3.28	2.71	2.71	15.67	4.36	4.36	100

All values are given in percent. *LICC*, linear intertemporal cross-correlation. *NICC*, nonlinear intertemporal cross correlation. *LCo*, linear cointegration. *NCo*, nonlinear cointegration. *LGC*, linear Granger casualty. *NGC*, nonlinear Granger causality. *SC*, shadow causality. *HC*, hidden causality.

2.5.2 The family of causalities: convergence or divergence?

Despite the fact that all the causality methods presented in this paper are normalized with values [-1, 1] or [0, 1], we cannot conduct a comparative analysis of them because they measure causality through different approaches (for a detailed understanding of what each method measures as causality see Section 2.2). Instead, what we can do is simply authenticate each individual method's reaction and probable predictive capacity to events of the financial market. To further justify the incomparability of the eight causality methods we undertake a similarity analysis (calculation of the percentage of similar links) for each of the 738 weeks of our evolutionary network and then calculate the average similarity across all weeks for all possible combinations of the causality methods. The results of this comparison are presented succinctly in Table 2.6. The highest average similarity between any two causality tools is 48.74% of links and occurs between *LICC* and *LGC*. The lowest average similarity between any two causality tools is 1.07% of links and occurs between *LICC* and *LCo*. Thus, we consider comparative analysis out of the question, at least in our experimental context. Maybe with the use of another filtering method, or even no filtering at all, we could find some common ground on which to carry out some comparisons.

2.5.3 Causal linkages: time-tested relationships

Having studied the evolution of the eight-causality metrics, and proved their dissimilarity, we move on to some distinct features, such as the most important links and assets. In Tables 2.7–2.11 we present the top ten links in terms of each causality separately, and the overall top thirty links in terms of all causalities together. As we can see from Table 2.7, in terms of *LICC* the most important causal relationship is 10Y US bond → 10Y German bond with *LICC* = 44:03%. However, in terms of *NICC*, the most causal pair is DAX 30 → CAC 40, with *NICC* = 72.49%. Furthermore, if we look at Table 2.8, we note that in terms of *LCo* the most important causal relationship is 10Y German bond → Bovespa, with LCD 79.40%, while in terms of *NC* the most causal pair is 10Y UK bond → 2Y US bond, with *NCo* = -28.99%. In terms of *LGC* the most important relationship is 3Y Italian bond → 2 to 3Y Spanish bond, with *LGC* = 56.91%. However, in terms of *NGC*, the most important causal pair is 10 Y Greek bond → 2 to 3Y Spanish bond, with *NGC* 45.25% (Table 2.9). In terms of SC the most important causal relationship is DAX 30 → BSE 30 Sensex with a score of 24.25%. Nevertheless, the most important causal relationship in terms of HC is 10Y US bond → 10Y German bond, with a

score of 48.64% (Table 2.10). Finally, if we consider the average ranking of all causalities, the most important causal pair overall is 10Y US bond → 2 to 3Y Spanish bond (Table 2.11).

2.5.4 Ranking financial assets in terms of causal influence

In order to rank the assets according to the causality they exert, we employ out strength centrality, which we average over the period for every causality method. The complete rankings can be seen in Tables 2.12–2.14. As we can see, the most causal asset in terms of *LICC* is the 10Y US bond, while that in terms of *NICC* is the Hang Seng index. Moreover, in terms of *LCo* the most influential asset is the 2Y Japanese bond, while in terms of *NCo* it is the DAX 30 index. *LGC* coincides with *LICC* in crowning the 10Y US bond the most causal, which is also the leading causal asset in terms of *NGC*. However, the results in terms of *SC* and *HC* are different, giving the BSE 30 Sensex index and the 2Y Japanese bond as the most causal assets, respectively. All in all, the most causal asset in terms of all causalities considered appears to be the 10Y US bond.

Furthermore, we note that, in agreement with Rahman and Mustafa (1997), *LCo* and *HC* unveil a "hidden" regime of causality occasionally monopolized by the bonds (see Tables 2.12 and 2.14). This result is astounding because *LCo* attests that those bonds exert the most powerful linear and profound long-term influence on the other assets and *HC* further reveals a nonlinear and consistent short-term causality exercised by those bonds.

The sovereign bond market may be seen as the reflection of a country's monetary policy. This might be a possible explanation of why the US bond is overall is the most influential one, given that the US economy is considered to be quite robust even when instabilities trigger.

The Chinese stock market as represented by the Shanghai Composite Index is overall the one market with the least influence. A possible reason could be the fact that China is isolated from foreign investors, compared especially to countries from Europe, North America and Australia. The key takeaway here, is that the movement of the Shanghai Composite Index Composite Index has little to no impact the performance of other stock market indices.

Table 2.7: Top 10 out of 600 links in terms of strength throughout the time period examined for linear inter temporal cross-correlation and nonlinear intertemporal cross-correlation.

Rank	***LICC***	**Score**	***NICC***	**Score**
1	10Y US bond → 10Y German bond	0.4403	DAX 30 → CAC 40	0.7249
1	10Y US bond → 10Y German bond	0.4403	DAX 30 → CAC 40	0.7249
2	10Y US bond → 10Y Australian bond	0.4105	S&P 500 → CAC 40	0.6951
3	10Y US bond → 10Y Japanese bond	0.4092	Nikkei 225 → CAC 40	0.5338
4	10Y US bond → 2to3Y Spanish bond	0.3956	S&P 500 → CAC 40	0.4241
5	DJIA → Nikkei 225	0.3834	CAC 40 → S&P 500	0.4065
6	Bovespa → BSE 30 Sensex	0.3766	10Y Italian bond → 3Y Italian bond	0.4010
7	10Y US bond → 10Y Swiss bond	0.3509	Hang Seng → CAC 40	0.3983
8	Bovespa → Hang Seng	0.3265	10Y German bond → 10Y UK bond	0.3807
9	S&P 500 → Nikkei 225	0.3075	Hang Seng → Bovespa	0.3726
10	DJIA → BSE 30 Sensex	0.3021	Hang Seng → S&P 500	0.3699

Causality symbols: x → y denotes that x influences y in the same direction, ie, past x increases cause future y increases (similarly for decreases), while x → y denotes that x influences y in the opposite direction, i.e. past x increases cause future y decreases and vice versa. Causality score: the number ascribed to each causal relationship is the average causal weight from x to y for the period January 4, 2000 to February 12, 2016 for the specific causality method.

Table 2.8: Top 10 out of 600 links in terms of strength throughout the time period examined for linear cointegration and nonlinear cointegration.

Rank	***LCo***	**Score**	***NCo***	**Score**
1	10Y German bond → Bovespa	0.7940	10Y UK bond → 2Y US bond	-0.2899
2	10Y UK bond → Bovespa	0.7899	Heng Seng → Bovespa	0.2547
3	2Y Australian bond → Bovespa	0.7493	DJIA → 3Y Italian bond	0.2520
4	10Y Swiss bond → Bovespa	0.7249	Oil → Shanghai Composite Index	0.2384
5	10Y Australian bond → Bovespa	0.7208	Nikkei 225 → 2Y US bond	-0.2384
6	10Y US bond → Bovespa	0.6937	10Y UK bond → 10Y US bond	-0.2303
7	2Y German bond → Bovespa	0.6924	Nikkei 225 → 2Y German bond	-0.2168
8	10Y Japanese bond → Nikkei 225	0.6815	10Y US bond → 2Y Japanese bond	0.2046
9	2Y US bond → Bovespa	0.6747	DAX 30 → 3Y Italian bond	0.2018
10	10Y Japanese bond → Bovespa	0.6667	10Y Swiss bond → 2to3Y Spanish bond	-0.1978

Causality symbols: x → y denotes that x influences y in the same direction, ie, past x increases cause future y increases (similarly for decreases), while x → y denotes that x influences y in the opposite direction, i.e. past x increases cause future y decreases and vice versa. Causality score: the number ascribed to each causal relationship is the average causal weight from x to y for the period January 4, 2000 to February 12, 2016 for the specific causality method.

Table 2.9: Top 10 out of 600 links in terms of strength throughout the time period examined for linear Granger causality and nonlinear Granger causality.

Rank	***LGC***	**Score**	***NGC***	**Score**
1	3Y Italian bond → 2to3Y Spanish bond	0.7940	10Y Greek bond → 2to3Y Spanish bond	-0.2899
2	10Y US bond → 2to3Y Spanish bond	0.7899	3Y Italian bond → 2to3Y Spanish bond	0.2547
3	10Y Greek bond → 2to3Y Spanish bond	0.7493	10Y UK bond → 2to3Y Spanish bond	0.2520
4	2Y German bond → 2to3Y Spanish bond	0.7249	10Y US bond → 2to3Y Spanish bond	0.2384
5	10Y UK bond → 2to3Y Spanish bond	0.7208	10Y US bond → 2to3Y Spanish bond	-0.2384
6	10Y Italian bond → 2to3Y Spanish bond	0.6937	10Y Italian bond → 2to3Y Spanish bond	-0.2303
7	10Y US bond → 10Y Australian bond	0.6924	2Y German bond → 2to3Y Spanish bond	-0.2168
8	2Y US bond → 2to3Y Spanish bond	0.6815	2Y US bond → 2to3Y Spanish bond	0.2046
9	DJIA → BSE 30 Sensex	0.6747	10Y US bond → 10Y Japanese bond	0.2018
10	10Y Australian bond → 2to3Y Spanish bond	0.6667	10Y Greek bond → 10Y German bond	-0.1978

Causality symbols: x → y denotes that x influences y in the same direction, ie, past x increases cause future y increases (similarly for decreases), while x → y denotes that x influences y in the opposite direction, i.e, past x increases cause future y decreases and vice versa. Causality score: the number ascribed to each causal relationship is the average causal weight from x to y for the period January 4, 2000 to February 12, 2016 for the specific causality method.

Table 2.10: Top 10 out of 600 links in terms of strength through out the time period examined for shadow causality and hidden causality.

Rank	***SC***	**Score**	***HC***	**Score**
1	DAX 30 → BSE 30 Sensex	0.7940	10Y US bond → 10Y German bond	-0.2899
2	BSE 30 Sensex → Bovespa	0.7899	10Y US bond → 10Y Australian bond	0.2547
3	BSE 30 Sensex → S&P 500	0.7493	10Y Italian bond → 10Y German bond	0.2520
4	BSE 30 Sensex → Oil	0.7249	2Y German bond → 2to3Y Spanish bond	0.2384
5	CAC 40 → BSE 30 Sensex	0.7208	10Y UK bond → 10Y German bond	-0.2384
6	Hang Seng → BSE 30 Sensex	0.6937	DJIA → S&P 500	-0.2303
7	BSE 30 Sensex → ASX200	0.6924	10Y Greek bond → 10Y Italian bond	-0.2168
8	Bovespa → BSE 30 Sensex	0.6815	2Y Japanese bond → 3Y Italian bond	0.2046
9	BSE 30 Sensex → Shanghai Composite Index	0.6747	10Y UK bond → 10Y Swiss bond	0.2018
10	DJIA → BSE 30 Sensex	0.6667	10Y German bond → 10Y Australian bond	-0.1978

Causality symbols: x → y denotes that x influences y in the same direction, ie, past x increases cause future y increases (similarly for decreases), while x → y denotes that x influences y in the opposite direction, ie, past x increases cause future y decreases and vice versa.

Table 2.11: Top 30 out of 600 links in terms of average strength across all causalities.

Rank	Causal relationship
1	10Y US bond → 2to3Y Spanish bond
2	10Y Greek bond → 2to3Y Spanish bond
3	3Y Italian bond → 2to3Y Spanish bond
4	10Y US bond → 10Y Australian bond
5	10Y UK bond → 2to3Y Spanish bond
6	10Y Italian bond → 2to3Y Spanish bond
7	2Y US bond → 2to3Y Spanish bond
8	2Y German bond → 2to3Y Spanish bond
9	10Y US bond → 10Y Japanese bond
10	DJIA → BSE 30 Sensex
11	S&P 500 → Hang Seng
12	10Y US bond → 10Y Swiss bond
13	10Y US bond → 10Y Swiss bond
14	DJIA → Nikkei 225
15	10Y Greek bond → 10Y German bond
16	10Y Australian bond → 2to3Y Spanish bond
17	Bovespa → BSE 30 Sensex
18	2Y Australian bond → 2to3Y Spanish bond
19	S&P 500 → Nikkei 225
20	Bovespa → Hang Seng
21	10Y German bond → Bovespa
22	10Y Swiss bond → Bovespa
23	DJIA → Hang Seng
24	DAX 30 → BSE 30 Sensex
25	10Y UK bond → Bovespa
26	2Y Australian bond → Bovespa
27	2Y US bond → 10Y Swiss bond
28	2Y German bond → Bovespa
29	Bovespa → Shanghai Composite Index
30	CAC 40 → BSE 30 Sensex

Causality symbols: x → y denotes that x influences y in the same direction, ie, past x increases cause future y increases (similarly for decreases), while x → y denotes that x influences y in the opposite direction, ie, past x increases cause future y decreases and vice versa. Causality score: the number ascribed to each causal relationship is the average causal weight from x to y for the period January 4, 2000 to February 12, 2016 for the specific causality method.

Table 2.12: Asset ranking in terms of out-strength centrality.

LICC		***NICC***		***LCo***	
Asset	**Strength**	**Asset**	**Strength**	**Asset**	**Strength**
10Y US bond	1.075	Hang Seng	0.871	2Y Japanese bond	1.377
DJIA	0.538	DAX 30	0.838	10Y Japanese bond	0.774
S&P 500	0.501	10Y Greek bond	0.830	10Y Swiss bond	0.529
10Y UK bond	0.501	10Y German bond	0.761	10Y UK bond	0.479
2Y US bond	0.486	S&P 500	0.713	10Y German bond	0.472
Bovespa	0.416	Nikkei 225	0.585	2Y German bond	0.464
10Y Greek bond	0.407	10Y Italian bond	0.521	10Y Greek bond	0.410
CAC 40	0.396	CAC 40	0.496	2Y Australian bond	0.393
DAX 30	0.391	10Y Swiss bond	0.383	10Y Australian bond	0.379
3Y Italian bond	0.223	2Y German bond	0.315	10Y US bond	0.377
2Y Australian bond	0.222	10Y US bond	0.302	10Y Italian bond	0.370

Oil	0.202	2to3Y Spanish bond	0.299	2Y US bond	0.314
10Y Japanese bond	0.185	10Y Australian bond	0.284	3Y Italian bond	0.311
10Y German bond	0.177	DJIA	0.282	2to3Y Spanish bond	0.163
BSE 30 Sensex	0.170	BSE 30 Sensex	0.282	Oil	0.021
2Y German bond	0.169	ASX 200	0.278	Shanghai Composite Index	0.000
Shanghai Composite Index	0.169	Bovespa	0.262	Bovespa	0.000
ASX 200	0.160	Oil	0.225	DJIA	0.000
10Y Swiss bond	0.145	2Y US bond	0.186	S&P 500	0.000
10Y Italian bond	0.142	10Y Japanese bond	0.181	DAX 30	0.000
10Y Australian bond	0.137	3Y Italian bond	0.138	Hang Seng	0.000
Hang Seng	0.122	Shanghai Composite Index	0.096	CAC 40	0.000
2Y Japanese bond	0.105	10Y UK bond	0.077	Nikkei 225	0.000
Nikkei 225	0.080	2Y Japanese bond	0.076	ASX 200	0.000
2to3Y Spanish bond	0.066	2Y Australian bond	0.059	BSE 30 Sensex	0.000

LICC, linear intertemporal cross-correlation. *NICC*, nonlinear intertemporal cross-correlation. *LCo*, linear cointegration. Score: out-strength centrality is calculated as the average for every node (asset) throughout the period January 4, 2000 to February 12, 2016 for the specific causality method.

Table 2.13: Asset ranking in terms of out-strength centrality.

NCo		***LGC***		***NGC***	
Asset	**Strength**	**Asset**	**Strength**	**Asset**	**Strength**
DAX 30	0.637	10Y US bond	2.429	10Y US bond	2.098
Nikkei 225	0.543	S&P 500	1.352	10Y Greek bond	1.480
10Y US bond	0.440	2Y US bond	1.267	10Y Swiss bond	1.242
10Y UK bond	0.440	DJIA	1.220	DJIA	1.195
Oil	0.410	Hang Seng	1.146	2Y US bond	1.193
3Y Italian bond	0.356	DAX 30	1.122	Bovepsa	1.093
CAC 40	0.324	Bovepsa	1.115	Hang Seng	1.024
2Y Australian bond	0.322	10Y UK bond	1.115	DAX 30	1.021
2Y German bond	0.319	3Y Italian bond	1.047	S&P 500	0.992
DJIA	0.306	10Y Greek bond	1.024	10Y UK bond	0.951
2Y US bond	0.276	2Y German bond	0.835	3Y Italian bond	0.940
BSE 30 Sensex	0.273	CAC 40	0.823	Oil	0.920
Hang Seng	0.263	10Y Japanese bond	0.821	10Y Italian bond	0.824
ASX 200	0.251	Shanghai Composite Index	0.811	2to3Y Spanish bond	0.821

Shanghai Composite Index	0.249	2Y Australian bond	0.786	ASX 200	0.815
S&P 500	0.245	10Y Italian bond	0.771	Nikkei 225	0.808
10Y German bond	0.234	Oil	0.746	10Y Japanese bond	0.708
10Y Australian bond	0.229	BSE 30 Sensex	0.729	CAC 40	0.699
10Y Greek bond	0.222	ASX 200	0.728	2Y German bond	0.698
10Y Swiss bond	0.210	10Y Australian bond	0.700	2Y Japanese bond	0.693
10Y Japanese bond	0.200	10Y Swiss bond	0.650	Shanghai Composite Index	0.685
10Y Italian bond	0.174	2Y Japanese bond	0.642	10Y Australian bond	0.678
Bovepsa	0.148	10Y German bond	0.512	BSE 30 Sensex	0.675
2to3Y Spanish bond	0.110	Nikkei 225	0.488	2Y Australian bond	0.521
2Y Japanese bond	0.107	2to3Y Spanish bond	0.460	10Y German bond	0.427

NCo, nonlinear cointegration. *LGC*, linear Granger causality. *NGC*, nonlinear Granger causality. Score: out-strength centrality is calculated as the average for every node (asset) throughout the period January 4, 2000 to February 12, 2016 for the specific causality method.

Table 2.14: Asset ranking in terms of out-strength centrality.

SC		***HC***		**Total (mean)**	
Asset	**Strength**	**Asset**	**Strength**	**Asset**	**Strength**
BSE 30 Sensex	1.150	2Y Japanese bond	0.423	10Y US bond	0.904
Shanghai Composite Index	0.486	10Y US bond	0.346	10Y Greek bond	0.630
10Y Greek bond	0.458	2Y German bond	0.257	DAX 30	0.542
S&P 500	0.399	10 UK bond	0.241	S&P 500	0.540
DJIA	0.342	2Y US bond	0.215	2Y US bond	0.516
2Y Japanese bond	0.323	10Y Greek bond	0.208	DJIA	0.507
Bovespa	0.307	DJIA	0.174	10Y UK bond	0.505
ASX 200	0.280	2Y Australian bond	0.156	2Y Japanese bond	0.468
2Y German bond	0.262	10Y Australian bond	0.152	Hang Seng	0.458
CAC 40	0.259	10Y Italian bond	0.140	Bovespa	0.430
10Y UK bond	0.234	S&P 500	0.121	10Y Swiss bond	0.422
DAX 30	0.225	10Y German bond	0.113	BSE 30 Sensex	0.415
10Y Italian bond	0.214	DAX 30	0.100	2Y German bond	0.415
Hang Seng	0.201	Bovespa	0.099	3Y Italian bond	0.409
Nikkei 225	0.200	Shanghai Composite Index	0.074	10Y Italian bond	0.395

10Y Australian bond	0.200	3Y Italian bond	0.068	10Y Japanese bond	0.384
Oil	0.195	10Y Swiss bond	0.048	CAC 40	0.379
3Y Italian bond	0.190	BSE 30 Sensex	0.040	10Y German bond	0.360
2Y US bond	0.188	Nikkei 225	0.039	10Y Australian bond	0.345
10Y German bond	0.184	Hang Seng	0.038	Oil	0.343
2Y Australian bond	0.179	CAC 40	0.034	Nikkei 225	0.343
10Y Japanese bond	0.176	ASX 400	0.031	2Y Australian bond	0.330
10Y Swiss bond	0.173	10Y Japanese bond	0.028	Shanghai Composite Index	0.321
2to3Y Spanish bond	0.173	Oil	0.024	ASX 200	0.318
2Y Japanese bond	0.165	2to3Y Spanish bond	0.020	2to3Y Spanish bond	0.264

SC, shadow causality. *HC*, hidden causality. Score: out-strength centrality is calculated as the average for every node (asset) throughout the period January 4, 2000 to February 12, 2016 for the specific causality method.

2.6 Network visualization

The extraordinary performance of the sovereign bonds led us to further examine our financial network's evolutionary behavior for each of the eight causality methods.

To that end, we plot four phases of the network for every causality method (see Figures 2.9–2.16). The phases record: (a) 2002 during the post-dot-com bubble burst, (b) 2008 during the global financial meltdown, (c) 2011 during the aftermath of the global crisis and (d) 2015 at the heart of the Chinese stock market crash.

On the onset of the *LICC* network (see Figure 2.9(a)), we can observe that the 2Y US bond is the predominant hub of causality, with the only competitive equity indexes being those of DAX30 and CAC40. During the global financial crisis, the 2Y US bond concedes its central role to the 10Y US bond, while the overall equity performance remained stable. Five years after the outbreak of the global crisis 10Y US bond still exerts the most causality on the network (see Figure 2.9(c)). However, its strength is diminished (see Figure 2.9(b)). As far as the equities are concerned, ASX 200 appears to be the most influential. Finally, during the Chinese stock market crisis it is really interesting to see that oil becomes a hub of causality.

Viewing the same financial network through the lens of *NICC* (see Figure 2.10), we witness a totally different situation: a disconnected network with substantially weaker relationships and no apparent hubs in all four phases. Unlike the *LICC* case, here we see little interaction between assets in different categories (equities and bonds). The only similarity appears to be the rising importance of oil as a key node during the Chinese stock market crash. Overall, we observe that *LICC* produces more stable relationships than *NICC*. However, this does not necessarily mean that *LICC* is better; it could mean that *LICC*, being a linear method, overestimated the causality intensity, while *NICC*, being a more "explorative" nonlinear method, is stricter in assigning higher scores. Investigation of the association between causality methods and their "meaning" is part of our future work.

We introduce *LCo* in Figure 2.11. Here we clearly see the reign of bonds throughout the four phases, with almost all of the linear long-term relationships being initiated by the bonds. Clearly, the equities subgroup is totally broken, with the individual equities being strongly influenced by bonds. Note that the 2Y Japanese bond is the hub of the financial network, particularly in parts (a) and (c) of Figure 2.11, where it is obvious that it influences most equities.

In the nonlinear form of the cointegration network (*NCo*) bonds still have a greater presence than equities (see Figure 2.12). During the aftermath of the dot-com bubble burst, the network appears quite mixed, with causal relationships between equities and bonds being formed interchangeably. BSE 30 Sensex appears to be the most influential equity, and the 2Y US bond the most central bond. Going into the global crisis, the network appears slightly more structured, with the leading asset role of the equities being handed over to DAX 30, and the strength of the 2Y US bond somewhat undermined.

After the crisis, the 2Y US bond is replaced by the 3Y Italian bond as the leading bond, and DAX 30 stands on par with it (see Figure 2.12 (c)). Finally, when the Chinese stock market crash takes place, no evident hub is observed. However, the subgroup of bonds is significantly more strongly connected than the subgroup of equities, which appear rather scattered.

LGC produces a financial network that, during the post-dot-com bubble burst (see Figure 2.13(a)), reveals a strong cluster of bonds and two broken groups of equities.

However, during the global financial crisis (see Figure 2.13(b)) the equities form a cluster and in fact outperform the group of disconnected bonds. DJIA, BSE 30 Sensex and Bovespa exert the most causality during that period. After the global financial crisis we see a strong presence of bonds, with dominance of the 10Y UK bond and the 2Y US bond. As far as the equities are concerned, the Hang Seng appears to be quite central (see Figure 2.13(c)). In the Chinese stock market crash, a recurring pattern of rising oil importance is visible (see Figure 2.13 (d)), while equities and bonds appear to share almost equally the causality in the network.

NGC appears to evolve in a parallel way (see Figure 2.14) to that of *LGC*: this is no surprise given that those processes are quite similar. The only significant difference seems to occur during the global financial crisis. Unlike the *LGC* case, here we can see a very strong cluster of bonds with immense centrality, and on the opposite side an equities subgroup divided into three, with DJIA the strongest index. Again, oil rises in significance during the Chinese stock market crash (see Figure 2.14 (d)).

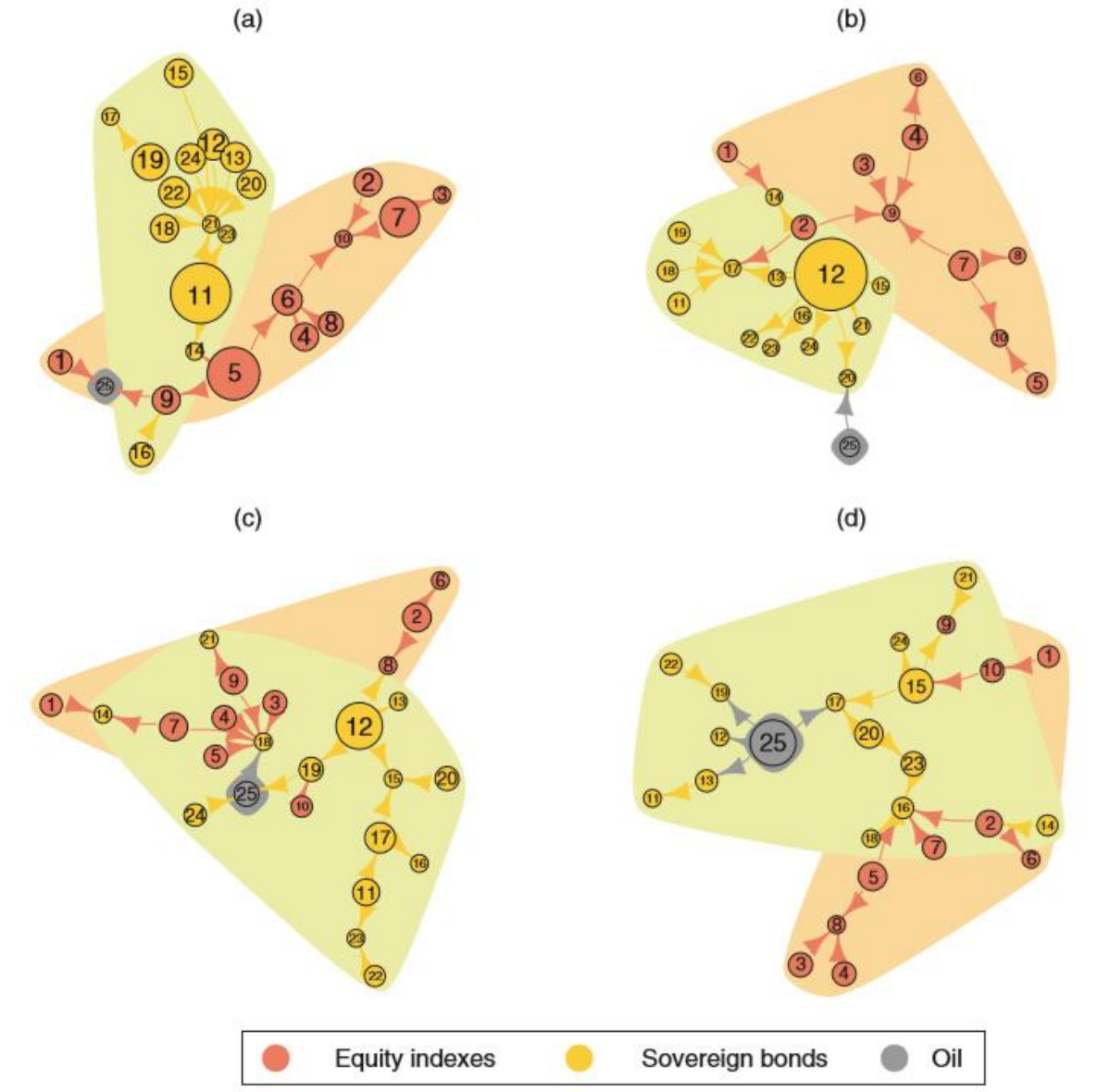
The network as seen through *SC* is quite intriguing (see Figure 2.15). During the post-dot-com bubble burst we observe an absolute balance between the equities and bonds, with oil lying in the middle of the network. When the global financial crisis breaks out, the 3Y Italian bond (see Figure 2.15(b)) polarizes the financial network and renders the equities group disconnected. In the aftermath of the crisis, the 10Y Greek bond stands as the hub of the financial network (Figure 2.15(c)). After four years, and into the Chinese stock market crash, the network appears quite clustered (see Figure 2.15(d)), with equities and bonds having few, trivial interactions.

The last building block in our analysis (*HC*) further confirms the existence of a bonds regime (see Figure 2.16). In the post-dot-com bubble burst, the Shanghai Composite Index index and a group of other equities are causally affected by the bonds cluster. The bonds cluster is mostly led by the 3Y Italian bond and the 10Y German bond (see Figure 2.16(a)). During the global financial crisis the group of equities seems further scattered, although the bonds seem

to be connected with even stronger causal relationships (see Figure 2.16 (b)). In the postmortem of the global financial crisis the 10Y and 2Y US bonds and the 10Y UK bond appear to concentrate most of the causal relationships. Ultimately, during the Chinese stock market crash we can see a network saturated absolutely by the bonds, with the 10Y German bond and the 10Y Greek bond being the undisputed hubs.

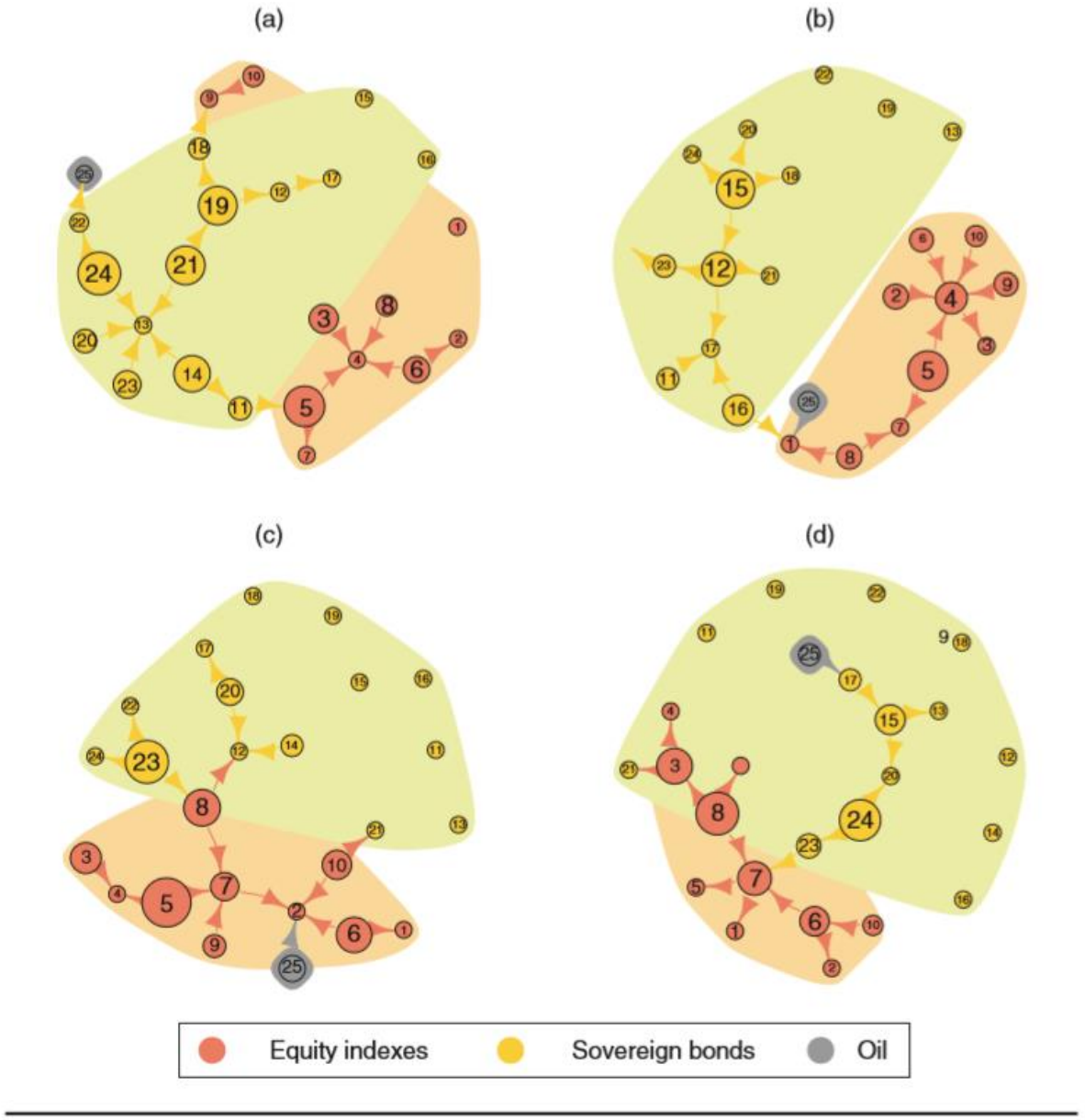
As a concluding remark here, the central role of oil during 2015 is a fact validated through many causality methods from our calculations. It is quite intriguing that this period, apart from the Chinese flash crash, also corresponds to the Arab Spring, a major political event in countries which are deeply entangled into the oil industry. Thus, there is a possibility that this central role of oil, in that period, somehow reflects that turbulence.

Figure 2.9: Linear intertemporal cross-correlation network (a) during the post-dot-com bubble burst, (b) during the global financial crisis, (c) after the global financial crisis and (d) during the Chinese stock market crash.



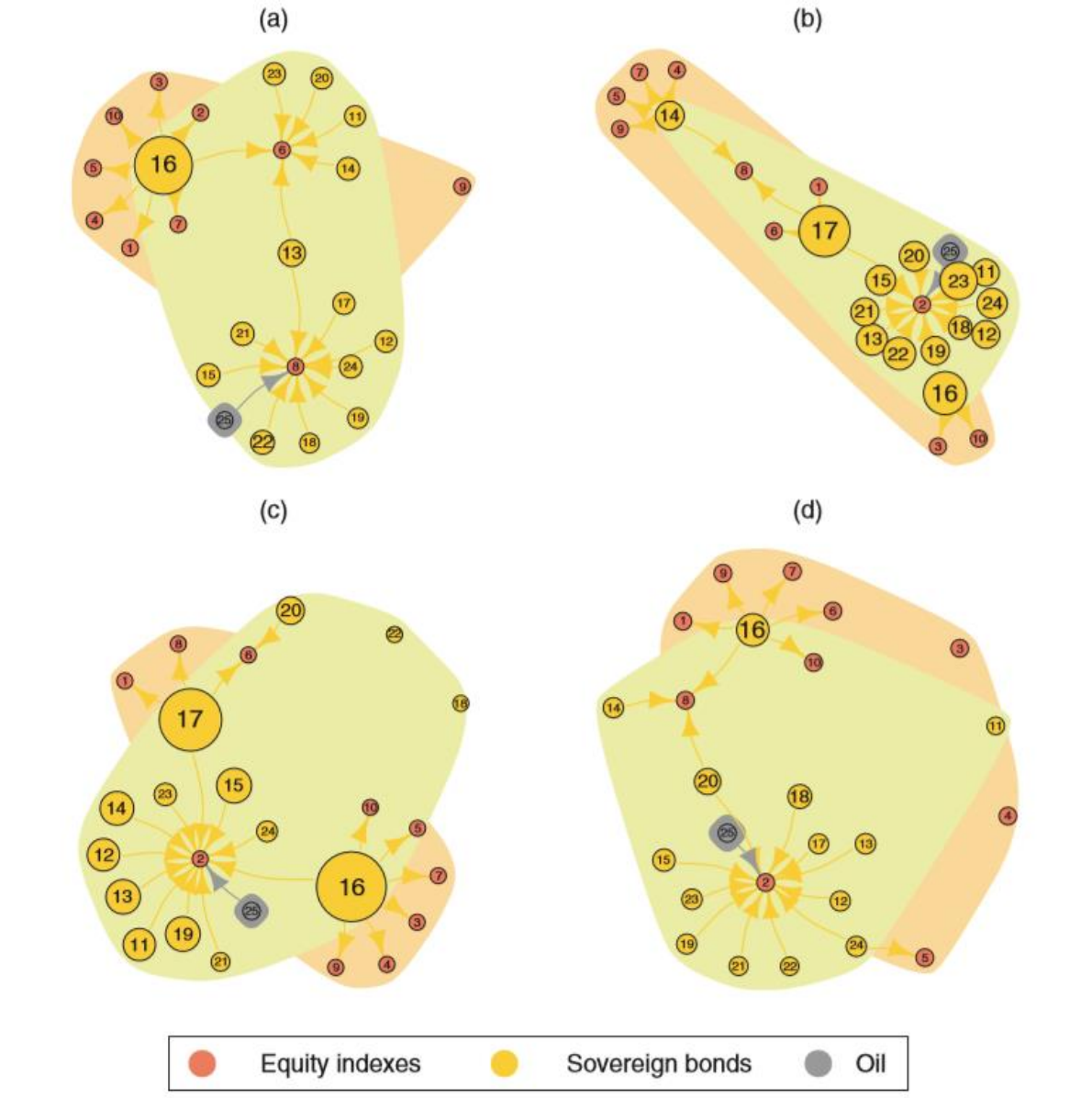
Node size: analogous to the node's out-strength centrality. Link width: analogous to the causality intensity. Link color: denotes the causality's origin (node category according to legend in each plot). Colored area: helps us understand visually the dominant asset category in terms of the network area (light orange-red for equities, light yellow-green for bonds and gray for oil).

Figure 2.10: Nonlinear intertemporal cross-correlation network (a) during the post-dotcom bubble burst, (b) during the global financial crisis, (c) after the global financial crisis and (d) during the Chinese stock market crash.



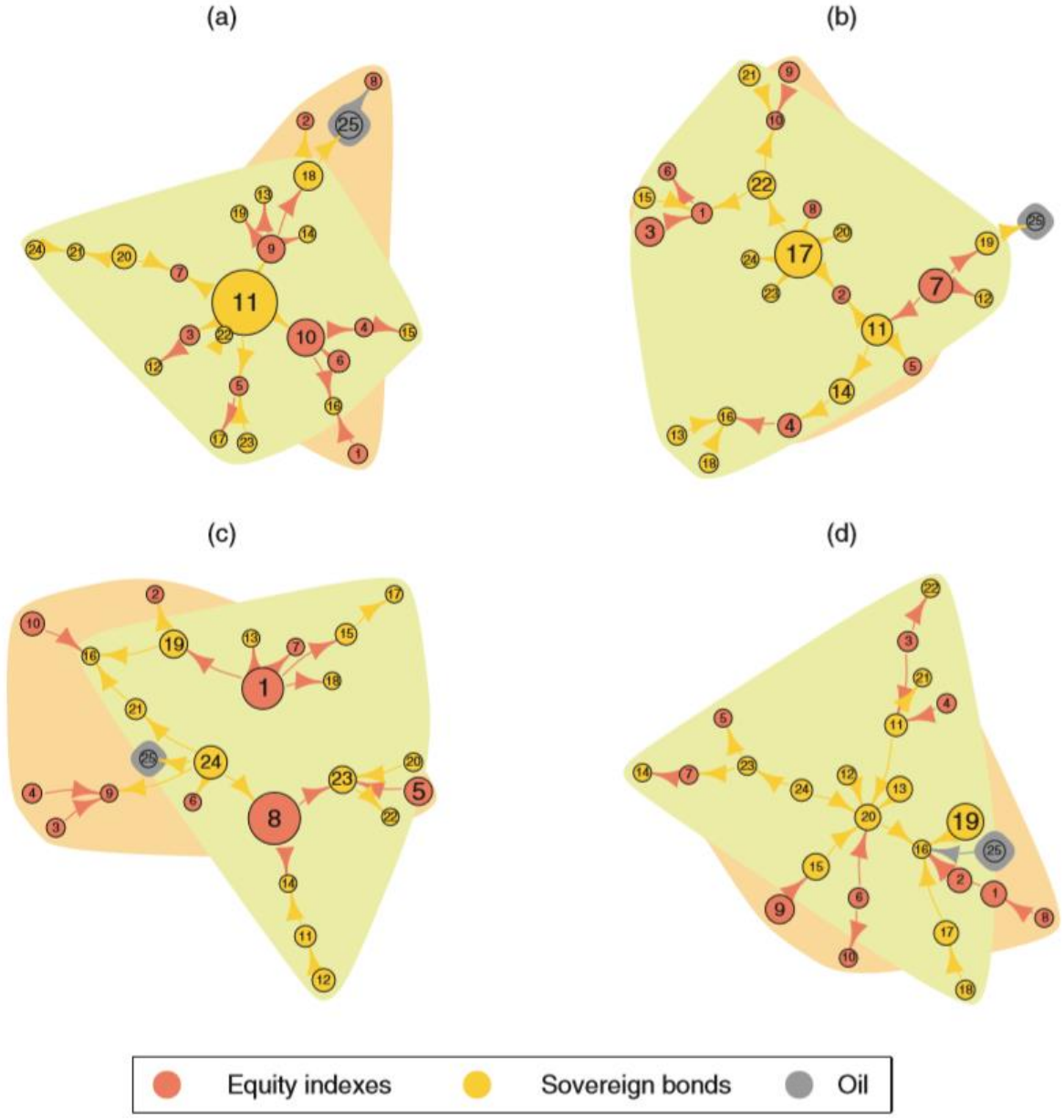
Node size: analogous to the node's out-strength centrality. Link width: analogous to the causality intensity. Link color: denotes the causality's origin (node category according to legend in each plot). Colored area: helps us understand visually the dominant asset category in terms of the network area (light orange-red for equities, light yellow-green for bonds and gray for oil).

Figure 2.11: Linear cointegration network (a) during the post-dot-com bubble burst, (b) during the global financial crisis, (c) after the global financial crisis and (d) during the Chinese stock market crash.



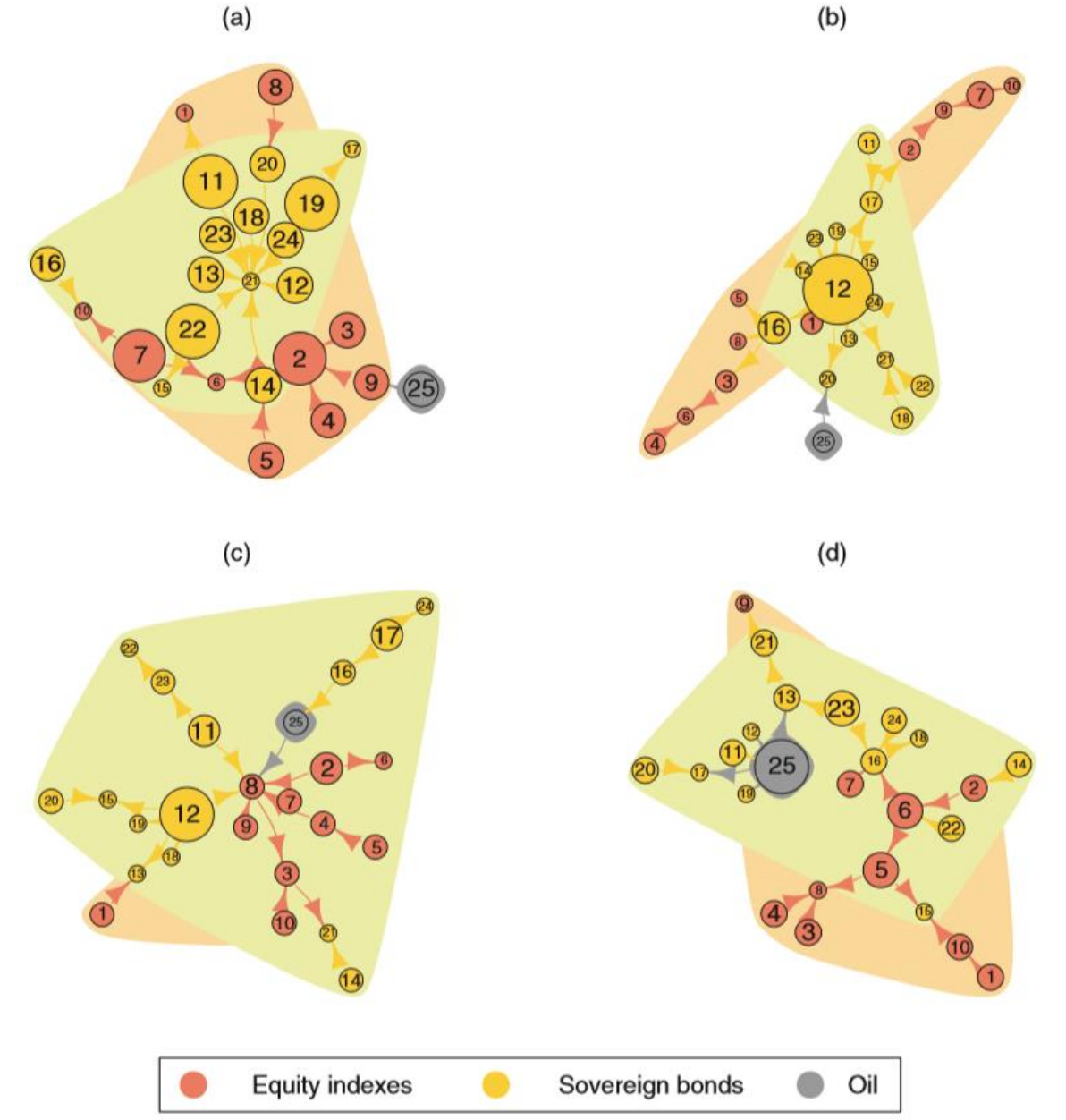
Node size: analogous to the node's out-strength centrality. Link width: analogous to the causality intensity. Link color: denotes the causality's origin (node category according to legend in each plot). Colored area: helps us understand visually the dominant asset category in terms of the network area (light orange-red for equities, light yellow-green for bonds and gray for oil).

Figure 2.12: Nonlinear cointegration network (a) during the post-dot-com bubble burst, (b) during the global financial crisis, (c) after the global financial crisis and (d) during the Chinese stock market crash.



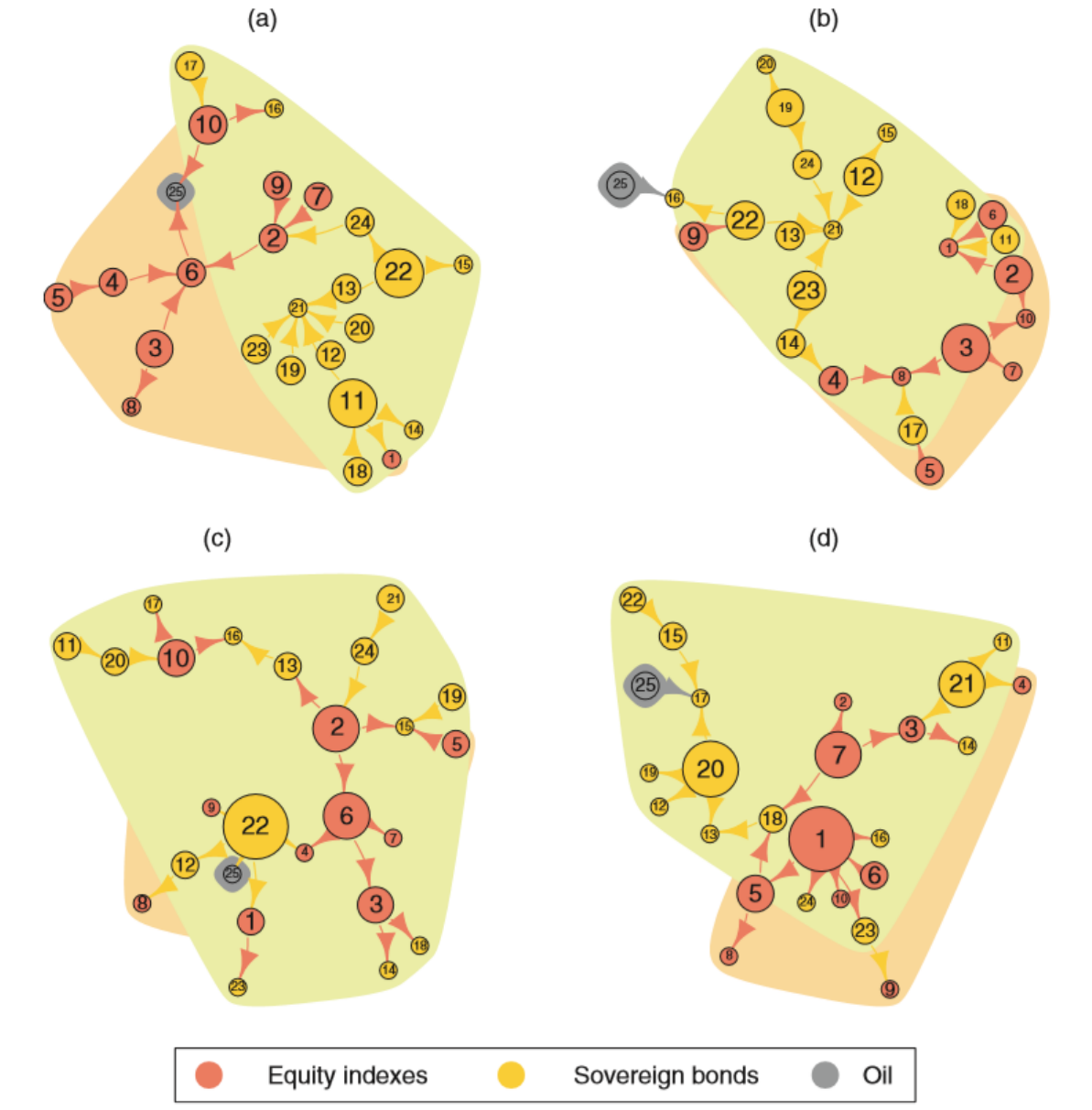
Node size: analogous to the node's out-strength centrality. Link width: analogous to the causality intensity. Link color: denotes the causality's origin (node category according to legend in each plot). Colored area: helps us understand visually the dominant asset category in terms of the network area (light orange-red for equities, light yellow-green for bonds and gray for oil).

Figure 2.13: Linear Granger causality network (a) during the post-dot-com bubble burst, (b) during the global financial crisis, (c) after the global financial crisis and (d) during the Chinese stock market crash.



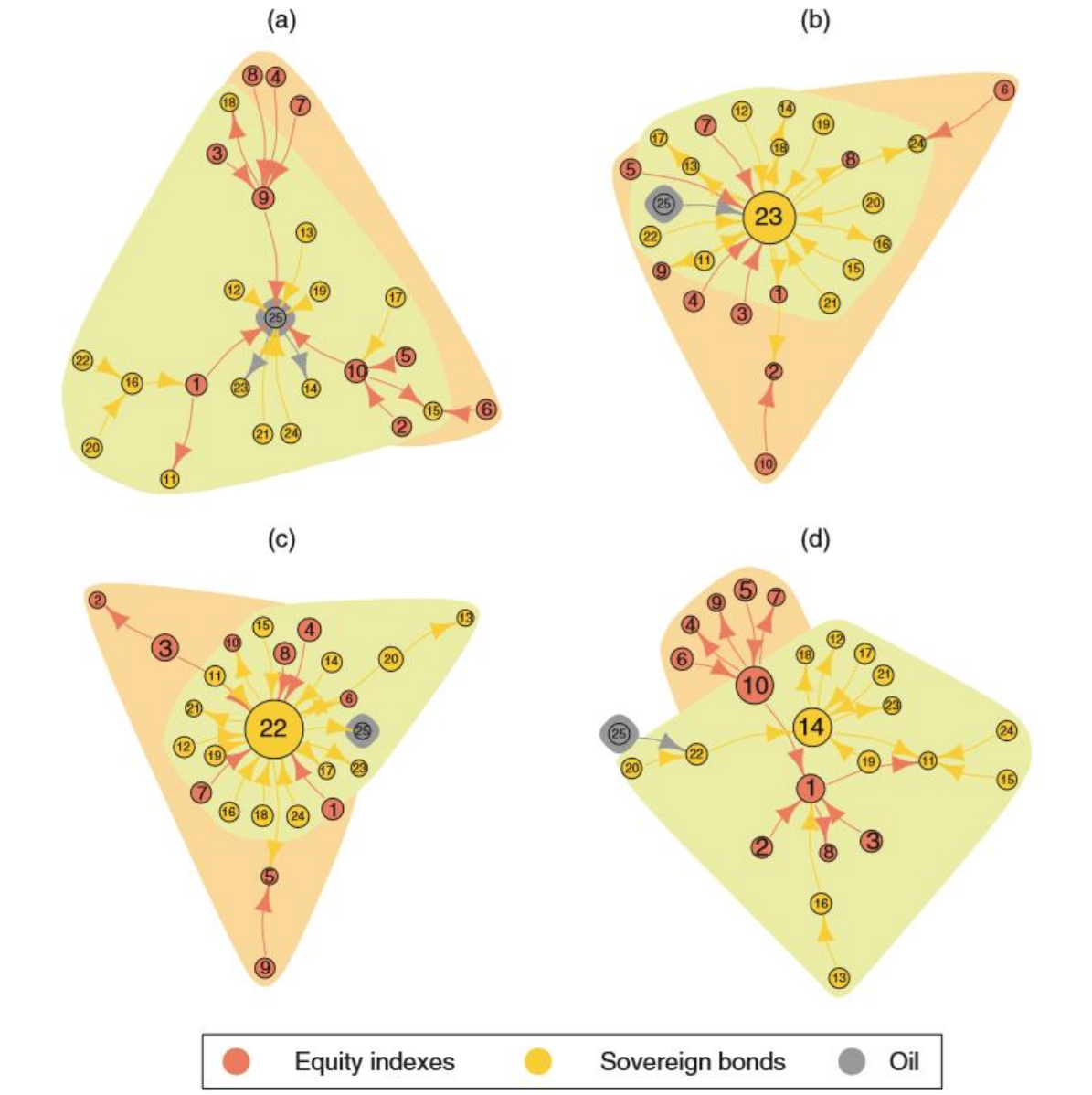
Node size: analogous to the node's out-strength centrality. Link width: analogous to the causality intensity. Link color: denotes the causality's origin (node category according to legend in each plot). Colored area: helps us understand visually the dominant asset category in terms of the network area (light orange-red for equities, light yellow-green for bonds and gray for oil).

Figure 2.14: Nonlinear Granger causality network (a) during the post-dot-com bubble burst, (b) during the global financial crisis, (c) after the global financial crisis and (d) during the Chinese stock market crash.



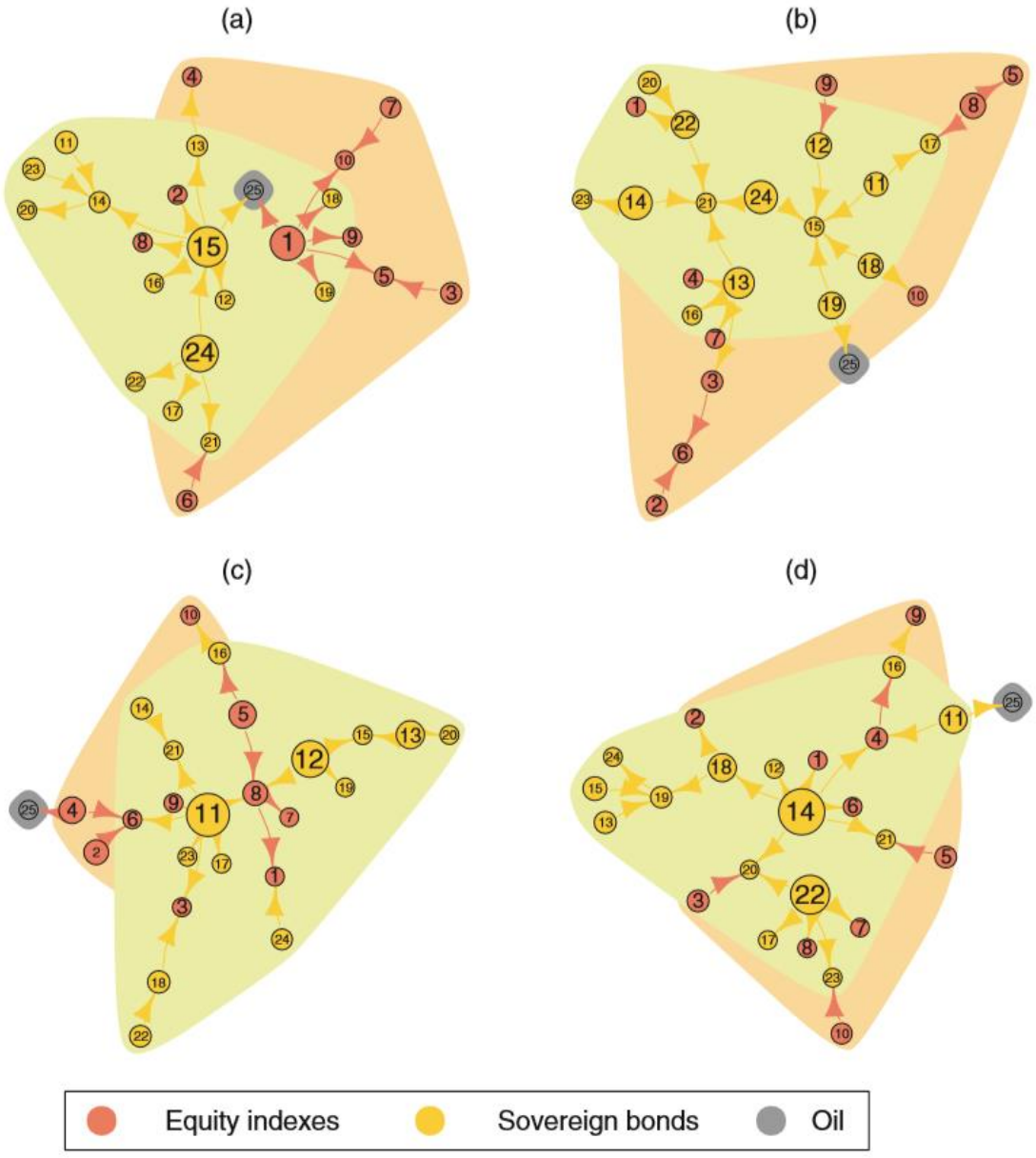
Node size: analogous to the node's out-strength centrality. Link width: analogous to the causality intensity. Link color: denotes the causality's origin (node category according to legend in each plot). Colored area: helps us understand visually the dominant asset category in terms of the network area (light orange-red for equities, light yellow-green for bonds and gray for oil).

Figure 2.15: Shadow causality network (a) during the post-dot-com bubble burst, (b) during the global financial crisis, (c) after the global financial crisis and (d) during the Chinese stock market crash.



Node size: analogous to the node's out-strength centrality. Link width: analogous to the causality intensity. Link color: denotes the causality's origin (node category according to legend in each plot). Colored area: helps us understand visually the dominant asset category in terms of the network area (light orange-red for equities, light yellow-green for bonds and gray for oil).

Figure 2.16: Hidden causality network (a) during the post-dot-com bubble burst, (b) during the global financial crisis, (c) after the global financial crisis and (d) during the Chinese stock market crash.



Node size: analogous to the node's out-strength centrality. Link width: analogous to the causality intensity. Link color: denotes the causality's origin (node category according to legend in each plot). Colored area: helps us understand visually the dominant asset category in terms of the network area (light orange-red for equities, light yellow-green for bonds and gray for oil).

2.7 Concluding remarks

Our results ascertain the existence of causal behavior among financial assets throughout the time period examined, with varying intensity according to the period under scrutiny. Such quantifiable causality among financial assets suggests that we can use past information in one financial asset to better anticipate the imminent performance of another asset. Given that efficient market hypothesis suggests that all information about an asset is immediately absorbed and thus reflected in the asset's current price, through causality analysis one might find a window of opportunity in the untapped information of assets about other assets. This outcome presents a potential unexplored range of trading strategies, whereby the institutional investor can use a holistic view (causality networks of assets) instead of single asset analysis in order to more optimally calibrate their portfolios.

We tested the percentage of similarity of common links between all causality methods and found that the most similar pair of causality-induced networks is on average less than 50% similar throughout the time period examined. Thus, we consider it meaning less to try to compare the results of different causality methods: every method deserves an explanation of its own. However, to the best of our knowledge, our study is the first attempt to unify various causalities for the production and analysis of financial networks.

We ranked the causal links as produced by each causality method and found that the most intense and protracted relationship across all causality methods is that of 10Y US bond → (ie, causing the prices of) 2to3Y Spanish bond. Furthermore, we ranked the financial assets in terms of averaged causality emanation, and uncovered a hidden "bonds regime", with the most causal asset being that of the 10Y US bond. Ultimately, we observed a recurring pattern of oil exerting increasing causality as the financial network entered the Chinese stock market crash.

3 Hidden Interactions in Financial Markets

The study of complex systems is hard-wired with the understanding of time series interdependencies. Specifically, financial markets have long been analysed as complex systems with the asset pricing being the foundation upon which structures such as financial networks are studied (Gai and Kapadia 2010; Preis et al. 2011; Gabaix et al. 2003; Boginski et al. 2005; Preis et al. 2010). Knowing the type of asset price interactions is vital to planning effective investment strategies or micro trading tactics (Markowitz 1952). Such knowledge is even more crucial if one considers the possibilities of switching regimes in financial markets (Preis et al. 2011).

A straightforward tool to gauge asset interactions is correlation but is known to be misleading in nonlinear systems such as stock markets (Yule 1926; Pearson 1895). The robust notion is that of causality (Pearl 2003) which has been approached via methods such as Granger causality (Granger 1969), cointegration (Engle and Granger 1987), transfer entropy (Schreiber 2000), convergent cross mapping (Sugihara et al. 2012) and S-maps (Deyle et al. 2016). Nevertheless, none of the aforementioned methods provide insight regarding more complex interactions, namely *dark causality*.

Our framework distinguishes between types of interactions (positive, negative and dark) with the full details are disclosed in section 3.5 and in SI Video[2]. In the case of positive causality X consistently causes the same patterns in Y, inversely in the case of negative causality X consistently causes the opposite patterns in Y. However, this is a small subset of the cases of normal interaction, as seen at the end of SI video and in Tables 3.1 and 3.2. In many cases X consistently causes patterns in Y, which are neither the same nor opposite. It is this third case, we observe complex causal interactions of temporal patterns that give rise to what we call dark causality, an obscure, yet substantial form of influence which is completely unapproachable through the lens of previous methods that gauge the nature of interactions (See Figures 3.5 to 3.8).

Positive causality (and positive correlation) is anticipated in cases of financial assets that are well-suited for asset pair trading (Gatev et al. 2006). Intuitively such assets cause same direction changes to one another. On the other hand, negative causality is expected in cases of

[2] SI Video: https://movie-usa.glencoesoftware.com/video/10.1073/pnas.1819449116/video-1

competing financial entities such as equities and bonds (Chow et al. 1999). The hidden nature of financial market interdependencies may be the missing link leading to a deeper understanding of market crashes and other emergent phenomena. This is why the emergence of dark causality is important, to uncover types of asset interactions which cannot be classified as purely positive or negative.

3.1 The nature of causality through contemporaneous patterns.

In complex systems that are comprised of deterministic or stochastic components, spatio-temporal dynamics are sculpted into a distinct attractor. According to dynamic systems theory when two time series X and Y are causally linked, they coexist in a common attractor, which is an embedded manifestation of their joint dynamic system. Consequently, each variable is imbued with information of the other's state (Takens 1981; Deyle and Sugihara 2011).

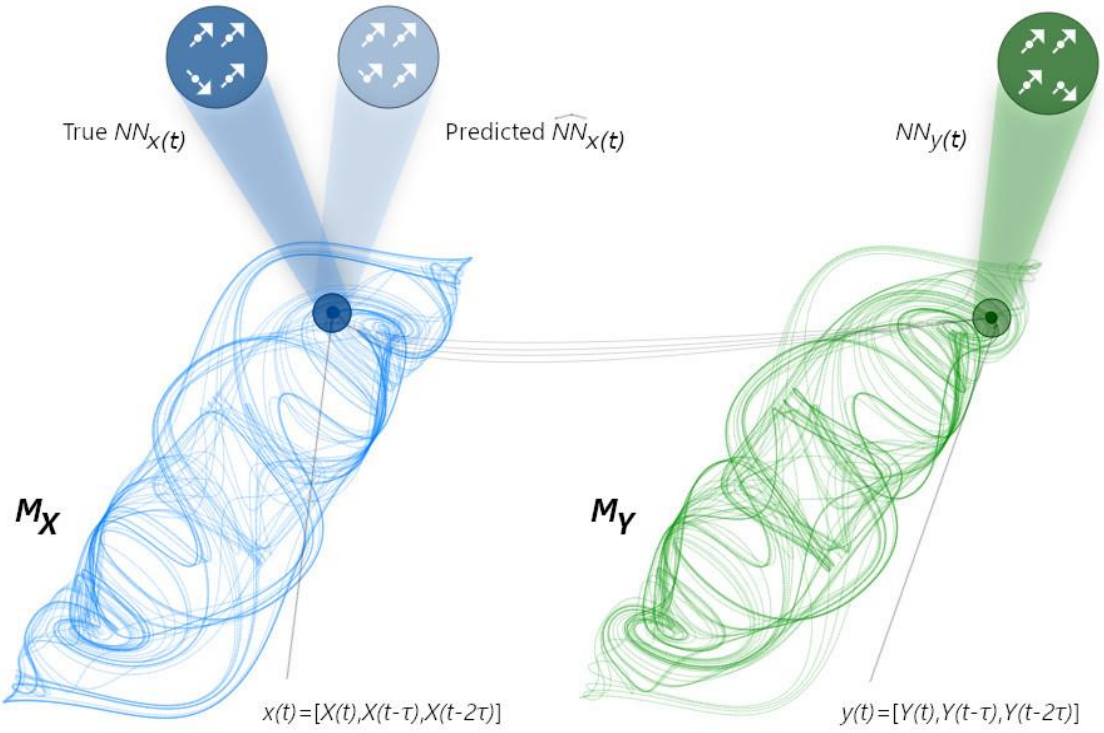
Nevertheless, the sharing of a common attractor is not sufficient to assess the nature of causality. Such intricate information is imprinted in the interplay of local spatio-temporal dynamics between the X's and Y's attractors, M_X and M_Y respectively, which is the focal point of our approach *Pattern Causality* (PC). Therefore, positive (or negative) causality from X to Y is manifested when the patterns in M_Y can accurately recall patterns of M_X and are of the same (or opposite) nature. When dark causality emerges, the patterns of M_X are imprinted and therefore recallable from M_Y, yet they are of neither the same nor opposite nature. This coupling of patterns is of complex nature, hence the naming convention “dark”.

Fundamentally, the idea is to see whether spatio-temporal neighborhoods of M_Y can consistently predict the patterns of counterpart neighbors of M_X. If the recalled patterns of M_X are correctly predicted, then X is causing Y to the extent of M_Y's predictive accuracy. Whether the nature of causality is positive, negative or dark depends on the correspondence of patterns between M_X and M_Y. Our approach is visually explained in Figure 3.1, below, with the full methodology recorded in section 3.5 and in SI Video.

PC is a method for detecting and quantifying the nature of causality which is based on two essential properties. Firstly, PC is characterized by consistency in its inference of causality's nature. In order for a causal relationship to be found as positive, negative or dark, the neighborhoods of M_Y have to systematically estimate correctly and in a consistent way the average patterns (signatures) in M_X. Thus, ephemeral correlations or evanescent causalities do not survive the PC trial. Consistency is what bestows PC with the ability to go beyond the

abstract quantification of causality. Secondly, the use of symbolic dynamics aids in the suppression of noise, thus making up for any expected and knowable noise in the system.

Figure 3.1: Pattern Causality (PC) (say from *X* to *Y*) is based on the accuracy of the memories about M_X's patterns which are embedded in M_Y's counterpart spatio-temporal neighbors. If *X* causes *Y*, the nearest neighbors of point *y(t)* ($NN_{y(t)}$) will correspond temporally to the nearest neighbors of x(t) ($NN_{x(t)}$). This enables us to "predict" patterns in the neighborhoods of M_X using patterns from M_Y. The more accurate our predictions are the higher the PC from *X* to *Y* is. To put it simply, if *X* causes *Y*, then patterns from M_X leave their "footprints" on M_Y, thus we can use patterns from M_Y to estimate the driving patterns from M_X.



3.2 Theoretical models

3.2.1 Positive causality (mutualism) model.

Positive causality suggests that two variables interact in such a way that changes in *X* cause consistently the same changes in *Y*. We use a mutualism model (see section 3.5, eq. 3.19) which describes two dynamically coupled variables with variable *X* exerting more influence on *Y* than vice versa (Fig. 3.2A). As we can see in Fig. 3.2D, PC detects the positive nature of

causality and is much higher from X to Y attesting to the asymmetric influence. Positive feedback loops can be described by such a system. These loops are very crucial to the understanding of many fields including: the dynamics of ecosystems (Crespi 2004), physiology for cardio-excitation – contraction coupling of the heart (Hall 2015) and also finance as indications of systemic risk (Arthur 1990), to name a few.

3.2.2 Negative causality (competition) model.

When negative causality is dominant this translates as X causing the opposite change in Y. For theoretical validation, we use a competition model (see section 3.5, eq. 3.20) of two dynamically coupled variables with X's impact being more intense than Y's (Fig. 3.2B). It is obvious from Fig. 3.2E that our approach correctly identifies the conflicting interaction and also reveals the asymmetry between X and Y. This case is also relevant to negative feedback loops, which are present in complex food webs (Neutel et al. 2007) and biological oxygen-dependent functions (Raven and Johnson 1999). Such loops are also encountered in the trading activity during financial crises and market crashes when investors are more pessimistic (Prechter 2001).

3.2.3 Dark causality (scapegoat) model.

For the case of dark causality, we employ a model (see section 3.5, eq. 3.21) which describes a scapegoat relationship (Fig. 3.2C). More specifically we simulate the interaction between two different prey population (e.g. lambs and rabbits) under the presence of a common predator (e.g. wolves). The design of our model allows the predator to hunt, at any given time unit, one type of prey exclusively. Each type of prey population may reproduce only when the other type is hunted in its place (e.g. when the wolves hunt lambs then the rabbits do reproduce). By calculating PC, we expose the hidden interaction between the two prey types (Fig. 3.2F), a relationship whose meaning is neither positive nor negative and falls into the category of dark causality. In finance, such relationships cannot be defined as beneficial or detrimental (as they can in ecology), but in the CDS example they are abundant and need to be further scrutinized.

3.3 Real Applications

Models exhibit almost ideal circumstances. This is when consistency is unhindered by observational error and system noise, which is not the case for shadow attractors reconstructed

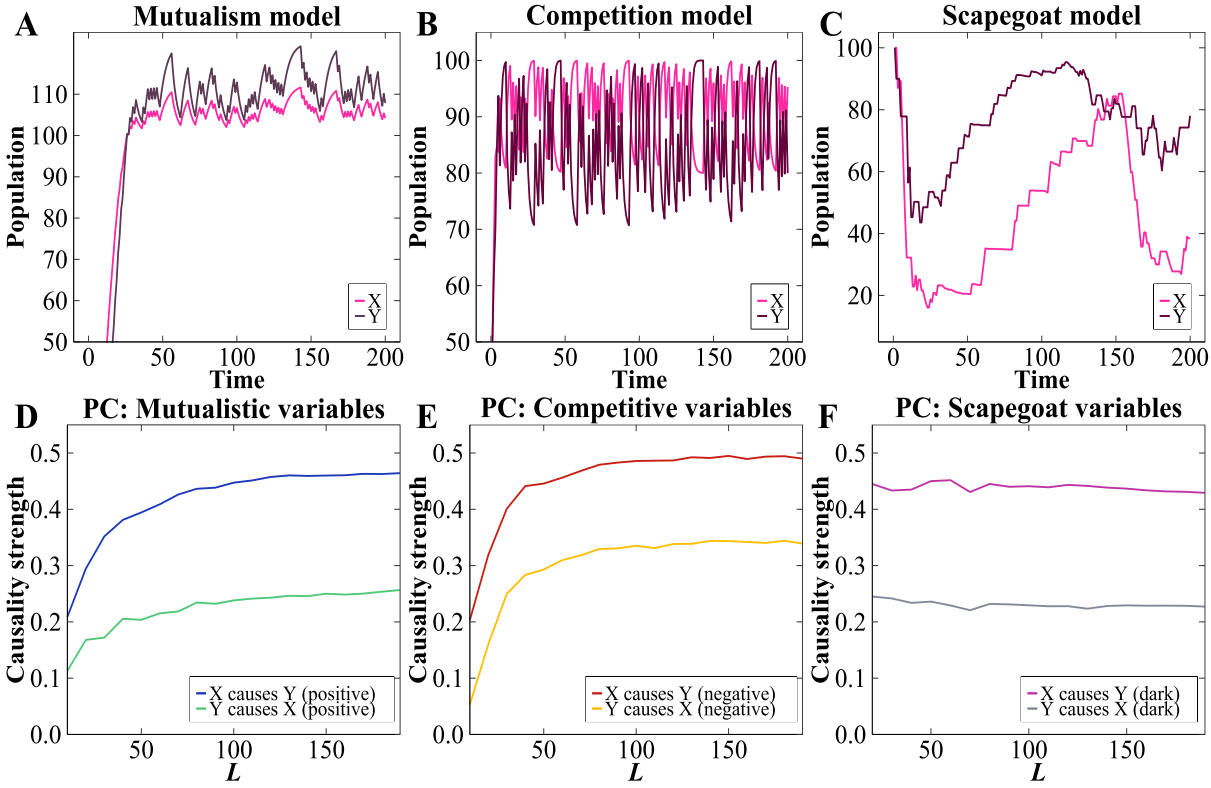
by real data. Nevertheless, even though noise and errors constrain the level of consistency, neighbors in shadow attractors can still recall significant amounts of spatio-temporal dynamics.

3.3.1 Pairs trading candidate assets.

Our first financial application of PC is on daily time series data of Apple (AAPL) and Microsoft (MSFT) retrieved from Datastream. The time span is from 1986-3-13 to 2018-8-6. The specific equities are chosen on the one hand for their popularity and on the other hand because they are usually studied in tandem (Da et al. 2011).

Causal interactions among these two assets (Fig. 3.3A) are distinctly positive (see Fig. 3.3B) which renders them ideal candidates for pairs trading, but ill-advised combination for a diversified portfolio strategy. Furthermore, the higher causative force from *MSFT* to *AAPL* suggests that the trading activity of *MSFT* is more influential on *AAPL* than vice versa, a fact that can aid in modeling forecasting.

Figure 3.2: Nature of causality in theoretical models. L is the time series (library) length. (A) Positive case: variables beneficial to each other (B) Negative case: variables competitive to each other (C) Dark case: variables involved in a persistent yet neither beneficial nor harmful relationship (D) PC between mutualistic variables. (E) PC between competitive variables. (F) PC in a scapegoat relationship. Color scheme: Blue and green are used for positive causality. Red and yellow are used for negative causality. Purple and grey are used for dark causality.

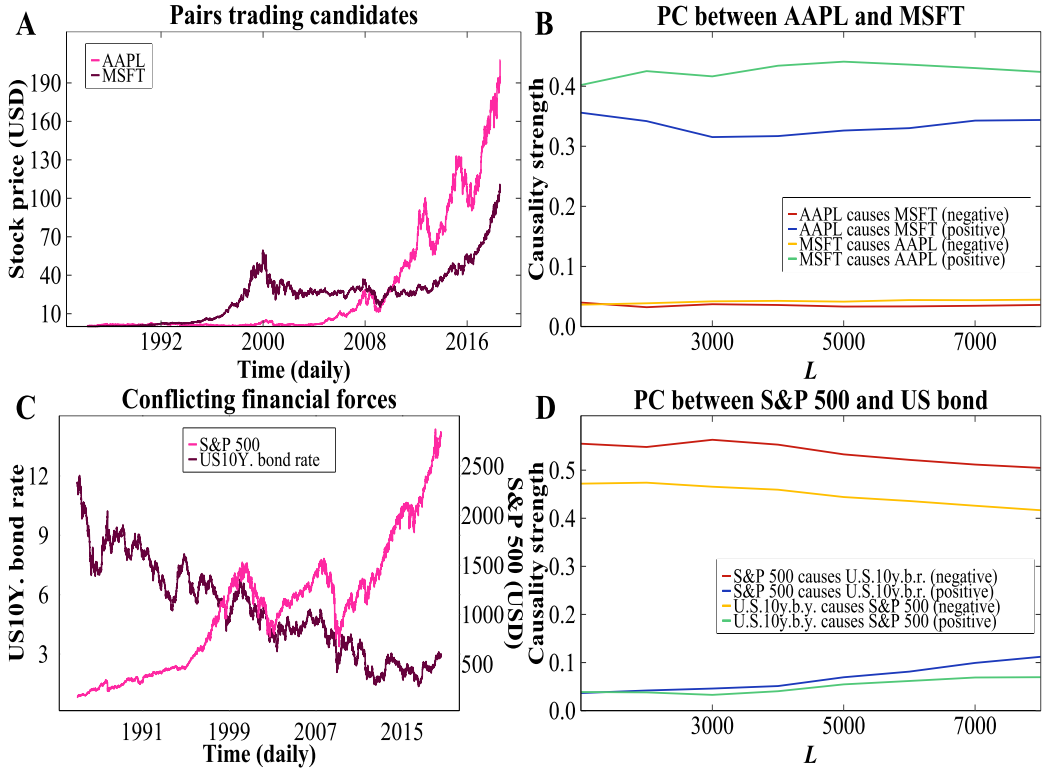


3.3.2 Conflicting financial forces.

Next, we apply PC in a classic example of opposing forces in finance (Kwan 1996) that of S&P 500 (as proxy of stock market performance) and U.S. government 10-year bond yield (Fig. 3.3C below). S&P 500 and bond yield data as available from Datastream. The time span is from 1985-1-2 to 2018-8-6.

The results in Fig. 3.3D validate the clasp between S&P 500 and U.S. 10-year bond, as negative causality, the previously assumed norm for decades. However, with a diminishing intensity since the year 2000. Contrary to the common view that government bond policy drives the stock market, our method supports the opposite. Policymakers may need to reconsider whether government bond yield leveraging affects the stock market or the assumption that they cause a negative feedback loop.

Figure 3.3: Nature of causality in financial data. L is the time series (library) length. (A) Daily time series of AAPL and MSFT. (B) PC reveals their positive causal interactions. (C) Daily time series of S&P 500 (right y axis) and US 10 years government bond yield (left y axis). (D) PC confirms their negative interaction. Color scheme: Blue and green are used for positive causality. Red and yellow are used for negative causality. Purple and grey are used for dark causality.



3.3.3 Dark causality in global sovereign CDS networks.

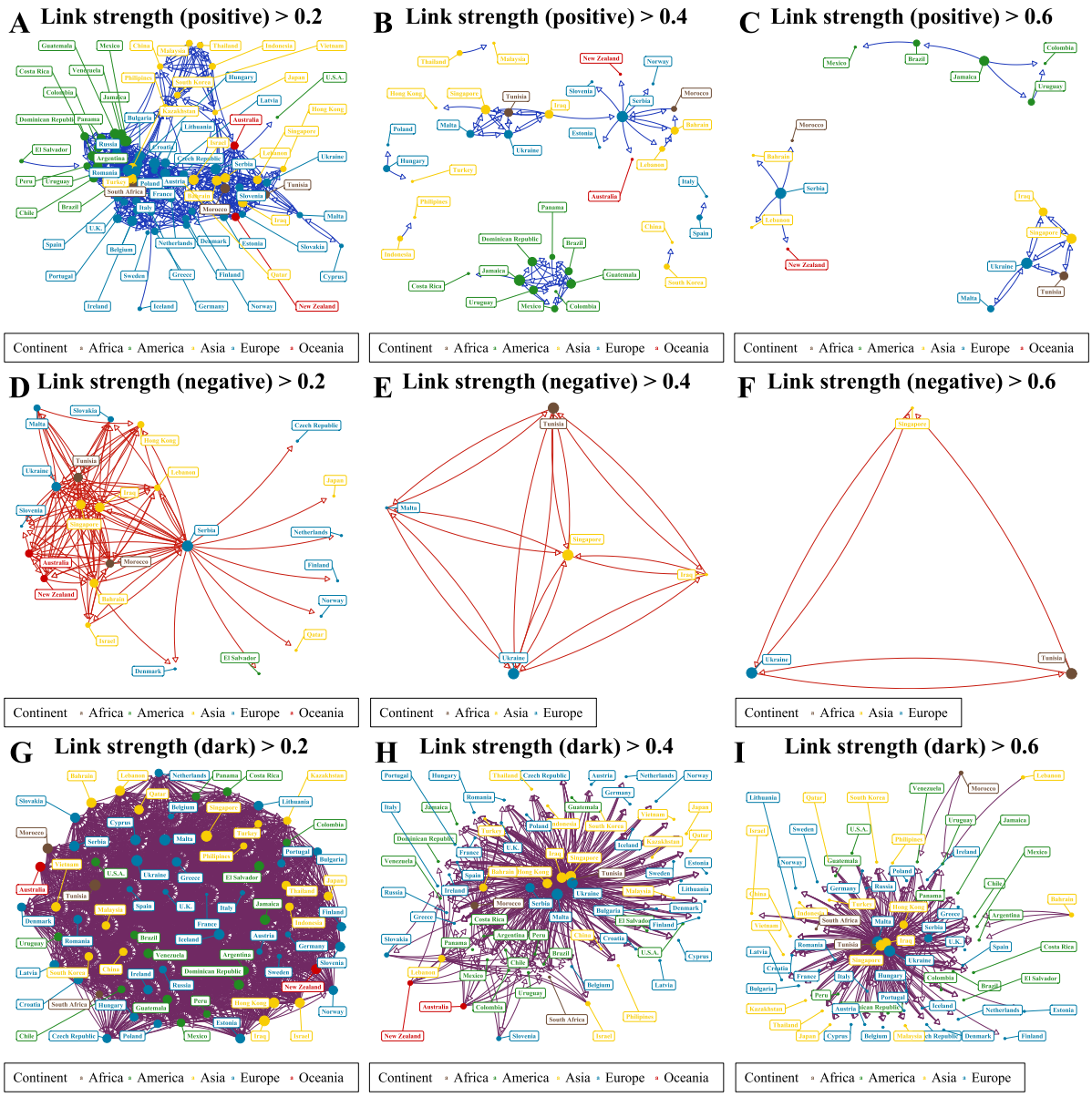
As a final and more complex system we analyze a dataset consisting of 69 sovereign CDS. The daily time series data were downloaded from Datastream and the time period spans from 2010-5-4 to 2018-8-6. CDS are relatively new derivates and academic research is still characterizing their mathematics and trying to understand their relation to financial crises (Gündüz and Kaya 2014) and determinants to the market (Peltonen et al. 2014).

Using PC as link weight we build the three emergent aspects of the CDS network (positive, negative and dark) (see Fig. 3.4A to 3.4I below). To get a broad view of the dominant nature of causality in this network we sequentially eliminate the weakest links from 0.2 link strength up to 0.6 by step of 0.2. Negative causality produces the most fragile network and when we eliminate links of up to 0.4 and 0.6 weight, the network is decimated down to 3 assets. The positive causality aspect of the network is slightly stronger with 15 assets remaining after the

final elimination process. On the contrary, the dark aspect of the causality network seems to be quite robust, since even after our final elimination step it remains with 65 assets and the connections remain rather dense.

The asymmetrical domination of dark causality means that the CDS market is strongly interconnected yet the correspondences in the shadow attractors are not clear about similar or opposite temporal patterns. This rise of complex interactions imprinted in the CDS attractors indicate non-trivial dynamics. Without a pure (positive or negative) form of causality the practice of few major dealers concentrating portfolios of large volumes of CDS (Peltonen et al. 2014) is not advisable in order to minimize systemic risk and credit risk exposure. In other words, the preeminence of dark causality can be seen as a sign of "*terra incognita*", (in Latin for unexplored land) in the already notorious CDS market (Acharya et al. 2009).

Figure 3.4: Nature of causality in CDS network.(A-C) Positive aspect of the network, gradual elimination of links below 0.2 (A), 0.4 (B) and 0.6 (C) PC. (D-F) Negative aspect of the network, gradual elimination of links below 0.2 (D), 0.4 (E) and 0.6 (F) PC. (G-I) Dark aspect of the network, gradual elimination of links below 0.2 (G), 0.4 (H) and 0.6 (I) PC. Overall dark causality is the most persistent type of PC. Link color scheme: Blue is used for positive causality. Red is used for negative causality. Purple is used for dark causality.



3.4 Concluding remarks

Greek philosophers such as Plato and Aristotle cerebrated the concept of causality (Hamilton et al. 2009; Hennig 2009). Through this study, we quantify the nature of causality among time series by gauging the correspondence of patterns in contemporaneously embedded neighborhoods and particularly the detection and quantification of dark causality. The more accurate the recalling ability of M_X's patterns about M_Y's patterns, the higher the causality from X to Y is. Whether the nature of causality is positive, negative or dark depends on the coupling of patterns between M_X and M_Y.

Causal networks are abundant in natural eco systems, biological processes and financial markets. More often than not, the mere quantification of causality is not enough. Species interact in complex and varied ways (e.g. symbiosis, competition or scapegoat relationship). Physiological functions are subject to underlying synergies which are not straightforward. The spectrum of causalities among financial assets is bountiful and their insights would be a great boom for economists and policymakers alike. By unveiling the innermost mechanics of dynamical systems, Pattern Causality offers a novel insight into the variety of causal interactions.

3.5 Algorithm details

3.5.1 Theoretical foundations of our method

In this section, we demonstrate the mathematical foundations of our treatment, and the details of extracting the nature of causality from time series. Our algorithm is primarily inspired by the theory of symbolic dynamics which was formally introduced by Morse and Hedlund (1938) and further adapted for time series by Bandt and Pompe (1993; 2002). Another important ingredient for our algorithm is dynamical systems theory (Alligood et al. 2000) and especially attractor reconstruction (Takens 1981).

Let us consider a discrete dynamical system that temporally evolves in an E-dimensional state space ($E \in \mathbb{N}$). Unless the system is completely stochastic, the orbits of its points will assemble into a d-dimensional attractor ($d_M \leq E$) $\boldsymbol{M}$. Let X be a state variable of the system that operates as a function which maps points from $\boldsymbol{M}$ to a real-valued scalar. Thus X can be measured through this mapping as a time series $\{X\}=\{X(1), \ldots, X(L)\}$ that records the orbits of points in $\boldsymbol{M}$. L is the time series length. By invoking the theory of time-delayed embedding,

the E time-lagged (with lag = $\tau \in \mathbb{N}$) values of $\{X\}$ spawn vectors ($x(t)=<X(t), X(t-\tau), ..., X(t-(E-1)\tau)>$) which can be used to create a diffeomorphically reconstructed attractor $\boldsymbol{M_X}$ of the original attractor $\boldsymbol{M}$ (Takens 1981; Deyle and Sugihara 2011; Josic 2000; Crutchfield 1979; Ruelle 1989). Furthermore, an intrinsic feature of delayed-coordinate embedding is that points $x(t)$ on $\boldsymbol{M_X}$ map 1:1 to points $m(t)$ on $\boldsymbol{M}$ and local neighborhoods on $\boldsymbol{M_X}$ map to local neighborhoods on $\boldsymbol{M}$ (Takens 1981; Deyle and Sugihara 2011).

Let us now consider another state variable of the system Y. Since X and Y originate from the same dynamical system they are dynamically coupled and as a consequence, contemporaneous neighborhoods on $\boldsymbol{M_X}$ and $\boldsymbol{M_Y}$ will map to each other (Sugihara et al. 2012; Takens 1981; Deyle and Sugihara 2011; Josic 2000; Crutchfield 1979; Ruelle 1989). PC focuses on the symbolic dynamics (patterns) of the neighborhoods in $\boldsymbol{M_X}$ and $\boldsymbol{M_Y}$, and examines how consistently, average patterns $\boldsymbol{P_X}$ (signatures) in local neighborhoods $\boldsymbol{NN_X}$ of $\boldsymbol{M_X}$ correspond to average patterns $\boldsymbol{P_Y}$ (signatures) in contemporaneous neighborhoods $\boldsymbol{NN_Y}$ of $\boldsymbol{M_Y}$.

3.5.2 Determining the nature of causality

To establish the nature of causality from a time series X to a time series Y (similarly from Y to X), first both $\boldsymbol{M_X}$ and $\boldsymbol{M_Y}$ are created from time-delayed vectors of X and Y. Then, for each point $y(t)$ in $\boldsymbol{M_Y}$ we extract the average pattern $P_{y(t)}$ from its nearest neighbors $NN_{y(t)}$ and from them we estimate the contemporaneous average pattern $\widehat{P_{x(t)}}$. The strength of causality is determined by the overall accuracy percentage between the estimated $\widehat{P_{x(t)}}$ and the actual $P_{x(t)}$. Regarding the nature of causality, if the correspondence is consistent between similar patterns ($\boldsymbol{P_X}$ similar to $\boldsymbol{P_Y}$) then positive causality defines the relationship of X and Y, whereas if opposite patterns are dominantly coupled (over similar ones) then negative causality is the case.

In our paradigm, the patterns we define for E=2 are: i) ↗: $X(t-\tau) < X(t)$, ii) ➡: $X(t-\tau) = X(t)$, iii) ↘: $X(t-\tau) > X(t)$ and these cover all the possible temporal patterns that characterize time series. Furthermore, for E=3 we define i) ↗↗: $X(t-2\tau) < X(t-\tau) < X(t)$, ii) ➡↗: $X(t-2\tau) = X(t-\tau) < X(t)$, iii) ↘↗: $X(t-2\tau) > X(t-\tau) < X(t)$, iv) ↗➡: $X(t-2\tau) < X(t-\tau) = X(t)$, v) ➡➡: $X(t-2\tau) = X(t-\tau) = X(t)$, vi) ↘➡: $X(t-2\tau) > X(t-\tau) = X(t)$, vii) ↗↘: $X(t-2\tau) < X(t-\tau) > X(t)$, viii) ➡↘: $X(t-2\tau) = X(t-\tau) > X(t)$, ix) ↘↘: $X(t-2\tau) > X(t-\tau) < X(t)$. The expansion on E=4 and beyond is plausibly derived in the same rationale.

There are cases however when persistent correspondence of patterns is confirmed through PC, yet the patterns are neither similar nor opposite (e.g. patterns of the for ↘↗ in *X* causing patterns of the form ↗↗ in *Y*), in such cases (which become abundant as *E* increases) we characterize this unclear nature of causality as dark, in the same manner that dark matter in the universe is a form of matter not fully understood yet abundant (Trimble 1987). For a complete blueprint on how to understand the nature of causality see Tables 3.1 and 3.2.

3.5.3 Remark on dark causality

Up until now scientific literature has concentrated on the obvious duality of interactions, i.e., positive/negative correlations (Pearson 1895), positively/negatively cointegrated time series (Engle and Granger 1987) and positive/negative nonlinear interactions (Deyle et al. 2016). However, through the lens of PC we discover a third form of interactions which classify as neither positive nor negative. For example, interactions in $E = 4$ such as pattern ↗↗↗↗ in *X* consistently causing pattern ↘↗↘↗ in *Y* can be described as: successive increases in *X* cause oscillations in *Y*. Such a form of causality from *X* to *Y* is neither positive nor negative, yet it is possible to exist.

3.5.4 Signature

In order to express a representation of the dominant dynamics in a spatio-temporal neighborhood on a given attractor ***M*** first we calculate the weighted average (see example below) of the patterns corresponding to the nearest neighbors ***NN***. The calculation of the weights in PC algorithm is done according to equation 3.3 below. Then we characterize as signature the pattern ***P*** which emerges from that weighted average.

For example, let us have four patterns:

- $s_1 =$ ↗↗ = (0.32, 0.45), with corresponding weight, $w_1 = 0.91$.
- $s_2 =$ ↘↗ = (-0.11, 0.51), with corresponding weight, $w_2 = 0.54$.
- $s_3 =$ ↗↗ = (0.13, 0.19), with corresponding weight, $w_3 = 0.82$.
- $s_4 =$ ↗↘ = (0.05, -0.08), with corresponding weight, $w_4 = 0.69$.

The weighted average in our example is:

$$S = \sum_{i=1}^{4} w_i s_i = 0.91 * (0.32, 0.45) + 0.54 * (-0.11, 0.51) + 0.82 * (0.13, 0.19) + 0.69 * (0.05, -0.08) = (0.3729, 0.7855)$$

Thus, the emergent average pattern is the signature of S:

$$\boldsymbol{P} = signature(S) = \nearrow\nearrow$$

3.5.5 PC Algorithm

Consider two time series of length L, $\{X\}=\{X(1), \ldots, X(L)\}$ and $\{Y\}=\{Y(1), \ldots, Y(L)\}$. Initially we derive an optimal combination of embedding dimension E and time delay τ. Optimal in our case would be with the least false nearest neighbors, given that our method relies heavily on the reliability of neighborhood information. For that purpose, we use the False First Nearest Neighbor algorithm (Krakovská et al. 2015) which calculates an optimal combination of both embedding dimension E and proper time delay τ simultaneously.

Then we retrieve the shadow attractors $\boldsymbol{M_X}$ and $\boldsymbol{M_Y}$ by using the lagged-coordinate vectors $x(t)=<X(t), X(t-\tau), ..., X(t-(E-1)\tau)>$ and $y(t)=<Y(t), Y(t-\tau), ..., Y(t-(E-1)\tau)>$ for $t = 1+(E-1)\tau$ to $t = L$. To calculate PC from X to Y (similarly from Y to X), for each point $y(t)$ in $\boldsymbol{M_Y}$ we find its $E+1$ nearest neighbors $NN_{y(t)}$, which is the minimum number of points needed for a bounded simplex in an E-dimensional space. From these $E+1$ nearest neighbors we need to keep three pieces of information: i) their time indexes $t_{y_1}, \ldots, t_{y_{E+1}}$, ii) their Manhattan (L1) distance from $y(t)$, and iii) their temporal patterns as described in the previous section. As a next step, we use the aforementioned pieces of information to estimate or "predict" the average contemporaneous pattern of $x(t)$. To do this first we calculate the average pattern (signature) $P_{y(t)}$:

$$P_{y(t)} = signature\left(S_{y(t)}\right),\ S_{y(t)} \in \mathbb{R}^E, \tag{3.1}$$

where

$$S_{y(t)} = \sum_{j=1}^{E+1} w_j^y s_j^y\ , w_j^y \in [0,1], s_j^y \in \mathbb{R}^E, \text{ for all } NN_{y(t)}, \tag{3.2}$$

$$w_j^y = \frac{e^{-d(y(t),y(t_j))}}{\sum_j e^{-d(y(t),y(t_j))}}, \ d\text{: Manhattan distance} \tag{3.3}$$

$$s_j^y = \left(\frac{y_j^{(2)} - y_j^{(1)}}{y_j^{(1)}}, \ldots, \frac{y_j^{(E+1)} - y_j^{(E)}}{y_j^{(E)}}\right), y_j \in \mathbb{R}, \tag{3.4}$$

$$y(t_j) = \left(Y(t_j), Y(t_j - \tau), \ldots, Y(t_j - (E-1)\tau)\right) = \left(y_j^{(1)}, \ldots, y_j^{(E+1)}\right), Y \in \mathbb{R}, \tag{3.5}$$

Then we estimate (i.e., predict) the mutual neighbors that correspond to $x(t)$ (contemporaneous to $y(t)$) by using the time indexes of $y(t)$'s nearest neighbors: $\widehat{NN_{x(t)}} = x_{t_{y_1}}, \ldots, x_{t_{y_{E+1}}}$. We calculate similarly the "predicted" average pattern $\widehat{P_{x(t)}}$ as follows:

$$\widehat{P_{x(t)}} = signature\left(\widehat{S_{x(t)}}\right), \widehat{S_{x(t)}} \in \mathbb{R}^E, \tag{3.6}$$

where

$$\widehat{S_{x(t)}} = \sum_{j=1}^{E+1} w_j^y s_j^{\hat{x}}, w_j^y \in [0,1], s_j^{\hat{x}} \in \mathbb{R}^E, \text{ for all } \widehat{NN_{x(t)}}, \tag{3.7}$$

$$s_j^{\hat{x}} = \left(\frac{\hat{x}_j^{(2)} - \hat{x}_j^{(1)}}{\hat{x}_j^{(1)}}, \ldots, \frac{\hat{x}_j^{(E+1)} - \hat{x}_j^{(E)}}{\hat{x}_j^{(E)}}\right), \hat{x}_j \in \mathbb{R}, \tag{3.8}$$

$$\hat{x}(t_j) = \left(\hat{X}(t_{y_1}), \hat{X}(t_{y_2}), \ldots, \hat{X}(t_{y_{E+1}})\right) = \left(\hat{x}_j^{(1)}, \ldots, \hat{x}_j^{(E+1)}\right), \hat{X} \in \mathbb{R}, \tag{3.9}$$

Finally, we calculate the real average pattern $P_{x(t)}$ from the actual nearest neighbors $\boldsymbol{NN_{x(t)}}$ of $x(t)$ to verify the estimation from Eq. 3.6:

$$P_{x(t)} = signature\left(S_{x(t)}\right), S_{x(t)} \in \mathbb{R}^E, \tag{3.10}$$

where

$$S_{x(t)} = \sum_{j=1}^{E+1} w_j^x s_j^x, w_j^x \in [0,1], s_j^x \in \mathbb{R}^E, \text{ for all } NN_{x(t)}, \tag{3.11}$$

$$w_j^x = \frac{e^{-d(x(t),x(t_j))}}{\sum_j e^{-d(x(t),x(t_j))}}, d\text{: Manhattan distance} \tag{3.12}$$

$$s_j^x = \left(\frac{x_j^{(2)} - x_j^{(1)}}{x_j^{(1)}}, \ldots, \frac{x_j^{(E+1)} - x_j^{(E)}}{x_j^{(E)}}\right), x_j \in \mathbb{R}, \tag{3.13}$$

$$x(t_j) = \left(X(t_j), X(t_j - \tau), \ldots, X(t_j - (E-1)\tau)\right) = \left(x_j^{(1)}, \ldots, x_j^{(E+1)}\right), X \in \mathbb{R}, \tag{3.14}$$

We repeat this procedure for every point in the shadow manifold $\boldsymbol{M_Y}$ and keep for possible pattern $\boldsymbol{P_Y}$ the weighted percentage of the occasions that the contemporaneous predicted dominant pattern $\widehat{P_X}$ equals the real dominant pattern $\boldsymbol{P_X}$. By this procedure, we fill in the PC pattern to pattern matrix (which is illustrated in Tables 3.1 and 3.2 for the case of $E = 1$ and $E = 2$, respectively):

$$PC[P_X, P_Y] = \sum_t \text{erf}\left(\frac{mean(|S_{y(t)}|)}{mean(|S_{x(t)}|)}\right), \text{erf: error squashing function.} \tag{3.15}$$

To determine the positive causality, we calculate the average accuracy regarding similar patterns ($\boldsymbol{P_Y}$ is the same as $\boldsymbol{P_X}$):

$$PC(Positive) = \frac{1}{L}\sum diag_{main}(PC), PC(Positive) \in [0, 1]. \tag{3.16}$$

Whereas for the extraction of negative causality, we calculate the average accuracy regarding opposite patterns ($\boldsymbol{P_Y}$ is the opposite of $\boldsymbol{P_X}$):

$$PC(Negative) = \frac{1}{L}\sum diag_{counter}(PC), PC(Negative) \in [0,1]. \quad (3.17)$$

Dark causality is the average accuracy regarding all the other combinations of patterns between $\boldsymbol{P_Y}$ and $\boldsymbol{P_X}$:

$$PC(Dark) = \frac{1}{L}\sum\left(PC \notin \left(diag_{main}(PC) \cup diag_{counter}(PC)\right)\right), PC(Negative) \in [0,1]. \quad (3.18)$$

3.6 Model dynamical systems and datasets of real applications

In this section, we provide details regarding the models and datasets in the main text. The models described below are based and modified and evolved from elementary Lotka-Volterra models studied in (Edelstein-Keshet 2004).

3.6.1 Positive causality system (Figure 3.2A)

Figure 3.2A demonstrates the phenomenon of ecological mutualism between two species derived from the following equations:

$$X(t+1) = 0.2X(t)[1 - X(t)/K_X + \varphi(p)0.1\,Y(t)/K_X], X \in \mathbb{R}^+, K_X \in \mathbb{R}^+.$$
$$Y(t+1) = 0.2Y(t)[1 - Y(t)/K_Y + \varphi(p)0.2\,X(t)/K_Y], Y \in \mathbb{R}^+, K_Y \in \mathbb{R}^+. \quad (3.19)$$

K_X and K_Y are the carrying capacities of species X and Y respectively, in our case both equal to 100. $\varphi(p)$ is a threshold function which is equal to 1 if p is greater or equal than 0.5 or 0 if otherwise, with p taking a random value between 0 and 1 for each t. The rationale behind $\varphi(p)$ is to make the model more realistic by having the two species not interacting in every turn. The starting conditions are: $X(1) = 10$, $Y(1) = 5$, and we use $E = 2$, $\tau = 1$, $L = 200$.

3.6.2 Negative causality system (Figure 3.2B)

Figure 3.2B demonstrates the phenomenon of ecological competition between two species derived from the following equations:

$$X(t+1) = 0.7X(t)[1 - X(t)/K_X - \varphi(p)0.2\,Y/K_X], X \in \mathbb{R}^+, K_X \in \mathbb{R}^+.$$
$$Y(t+1) = 0.7Y(t)[1 - Y(t)/K_Y + \varphi(p)0.3\,X/K_Y], Y \in \mathbb{R}^+, K_Y \in \mathbb{R}^+. \quad (3.20)$$

K_X and K_Y are the carrying capacities of species X and Y respectively, in our case both equal to 100. $\varphi(p)$ is a threshold function which is equal to 1 if p is greater or equal than 0.5 or 0 if otherwise, with p taking a random value between 0 and 1 for each t. The rationale behind $\varphi(p)$

is to make the model more realistic by having the two species not interacting in every turn. The starting conditions are: *X(1)=50, Y(1)=50* and we use *E = 2, τ = 1, L = 200.*

3.6.3 Dark causality system (Figure 3.2C)

Figure 3.2C displays the relationship between two species of prey *X* and *Y* that act as scapegoat to each other under the presence of a common predator *Z*. Such a system is described by the following equations:

$$X(t+1) = -\varphi_{XZ}(p)0.002X(t)Z(t) + \varphi_X(p)0.2X(t)\left(K_X - X(t)\right)/K_X, X \in \mathbb{R}^+, K_X \in \mathbb{R}^+.$$

$$Y(t+1) = -\varphi_{YZ}(p)0.002Y(t)Z(t) + \varphi_Y(p)0.2Y(t)\left(K_Y - Y(t)\right)/K_Y, Y \in \mathbb{R}^+, K_Y \in \mathbb{R}^+. \quad (3.21)$$

$$Z(t+1) = -0.2Z(t) + \varphi_{XZ}(p)0.005X(t)Z(t) + \varphi_{YZ}(p)0.010Y(t)Z(t), Z \in \mathbb{R}^+, K_Z \in \mathbb{R}^+.$$

K_X *and* K_Y are the carrying capacities of species *X* and *Y* respectively, in our case both equal to 100. $\varphi_{XZ}(p)$, $\varphi_X(p)$, $\varphi_{YZ}(p)$ and $\varphi_Y(p)$ are threshold functions.

- $\varphi_{XZ}(p)$ is equal to 1 if *p* is between 0.666 and 1 or 0 if otherwise
- $\varphi_X(p)$ is equal to 1 if *p* is between 0.166 and 0.333 or 0 if otherwise
- $\varphi_{YZ}(p)$ is equal to 1 if *p* is between 0.333 and 0.666 or 0 if otherwise
- $\varphi_Y(p)$ is equal to 1 if *p* is between 0 and 0.166 or 0 if otherwise

with *p* taking a random value between 0 and 1 for each *t*. The rationale behind $\varphi_{XZ}(p)$, $\varphi_X(p)$, $\varphi_{YZ}(p)$ and $\varphi_Y(p)$ is to make the model more realistic by having the species not interacting in every turn. Besides a prey species (say *X*) is more likely to breed when the other prey species (say *Y*) are hunted in their place, as scapegoats, by the common predator. The starting conditions are: *X(1)=100, Y(1)=100, Z(1)=50* and we use *E = 2, τ = 1, L = 200.*

3.6.4 A pair of equities (Figures 3.3A-B)

In Figures 3.3A-B we demonstrate the application of PC on daily time series data of Apple (AAPL) and Microsoft (MSFT) equities. The data are retrieved via Thomson Reuters Datastream. The time span is from 1986-3-13 to 2018-8-6. The parameters we use are $E = 3$, $\tau = 1$, $L = 8000$.

3.6.5 Stock market performance versus government bond yield (Figures 3.3C-D)

In Figures 3.3C-D we demonstrate the application of PC to the pair of S&P 500 (as proxy of stock market performance) and U.S. government 10-year bond yield. The data are retrieved via Thomson Reuters Datastream. The time span is from 1985-1-2 to 2018-8-6. The parameters we use are $E = 3$, $\tau = 1$, $L = 8000$.

3.6.6 Complex financial network of sovereign CDS (Figures 3.4A-I)

The sovereign CDS data we used to analyze with PC are from the following countries: Argentina, Tunisia, Venezuela, Czech Republic, Dominican Republic, Germany, Brazil, France, Greece, Hong Kong, Ireland, Jamaica, Japan, Bahrain, Belgium, Denmark, Norway, Spain, Sweden, Thailand, Netherlands, Lebanon, Malaysia, New Zealand, Uruguay, China, Austria, Bulgaria, Chile, Colombia, Costa Rica, Croatia, Cyprus, El Salvador, Estonia, Guatemala, Hungary, Iceland, Indonesia, Iraq, Italy, Kazakhstan, Latvia, Lithuania, Malta, Panama, Peru, Poland, Portugal, Serbia, Singapore, Slovenia, South Africa, Philippines, Turkey, Romania, Russia, Slovakia, Vietnam, Israel, Qatar, Ukraine, U.K., Mexico, Finland, South Korea, Morocco, U.S.A., Australia. The data are retrieved via Thomson Reuters Datastream. The time span is from 2010-5-4 to 2018-8-6. The parameters we use are: $E = 3$, $\tau = 1$, $L = 1000$.

3.6.7 Predator-prey interactions in the Didinium-Paramecium system (Figures 3.9A-B)

Next, we apply PC on experimental time series from the classical Didinium-Paramecium system. For a thorough description of the experimental conditions refer to the supplementary material of (Sugihara et al. 2012). The data analyzed can be found at http://robjhyndman.com/tsdldata/data/veilleux.dat. The first 10 data points were removed to eliminate transient behavior in the initial period of the experiment. The parameters we use are: $E = 3$, $\tau = 1$. To account for the limited length of time series, we used leave-one-out cross validation for the PC analysis.

3.6.8 Complex ecosystem of Sardine-Anchovy-SST (Figures 3.9C-F)

This dataset was also studied by (Sugihara et al. 2012) and details are also found in their supplementary material. Sea surface temperature (SST) data are available from: http://shorestation.ucsd.edu/active/index_active.html). Landings data for sardine and anchovy were

taken from two sources: 1. (1928-2002) NOAA Southwest Fisheries Science Center (http://las.pfeg.noaa.gov:8080/las_fish1/servlets/dataset?catitem=2) 2. (2003-2006) California Department of Fish and Game (http://www.dfg.ca.gov/marine/landings05.asp). The parameters we use are $E = 3$, $\tau = 1$. To address the limited length of time series, we used leave-one-out cross validation was used for the analysis.

3.6.9 Complex physiological system of heart-lungs-blood oxygen concentration (Figures 3.10A-F)

This dataset was used in the Santa Fe time series competition as Dataset B and was also studied by (Schreiber 2000). The dataset consists of heart rate, breath rate and blood oxygen concentration time series from a patient in the Beth Israel Deaconess Medical Center in Boston, Massachusetts (Rigney et al. 1993). Details about the dataset and very dataset can be found in https://www.physionet.org/physiobank/database/santa-fe/. The parameters we use are $E = 3$, $\tau = 1$, $L = 1000$.

Figure 3.5: Pearson cross correlation for the three models of positive (Eq. 3.19), negative (Eq. 3.20) and dark causality (Eq. 3.21). As we can see correlation is unstable for the case of competitive variables (Fig. 3.1B) and is completely irrelevant (by design) for the case of dark causality (Fig. 3.1C) by taking positive fluctuating values.

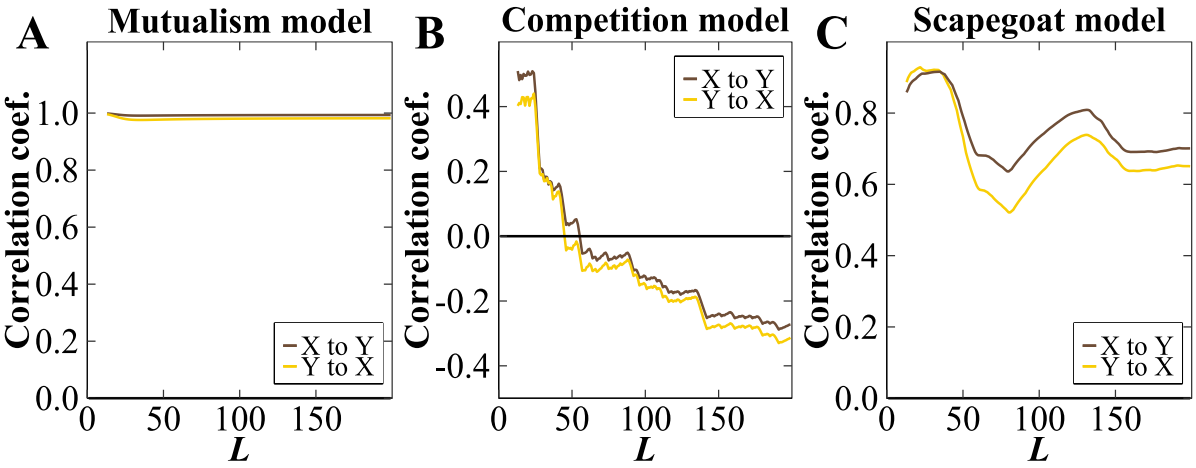


Figure 3.6: Beta coefficient of cointegration for the three models of positive (Eq. 3.19), negative (Eq. 3.20) and dark causality (Eq. 3.21). The coefficient of cointegration is unstable for the case of competitive variables (Fig. 3.2B) and cannot by default uncover any meaningful result for the case of dark causality (Fig. 3.2C).

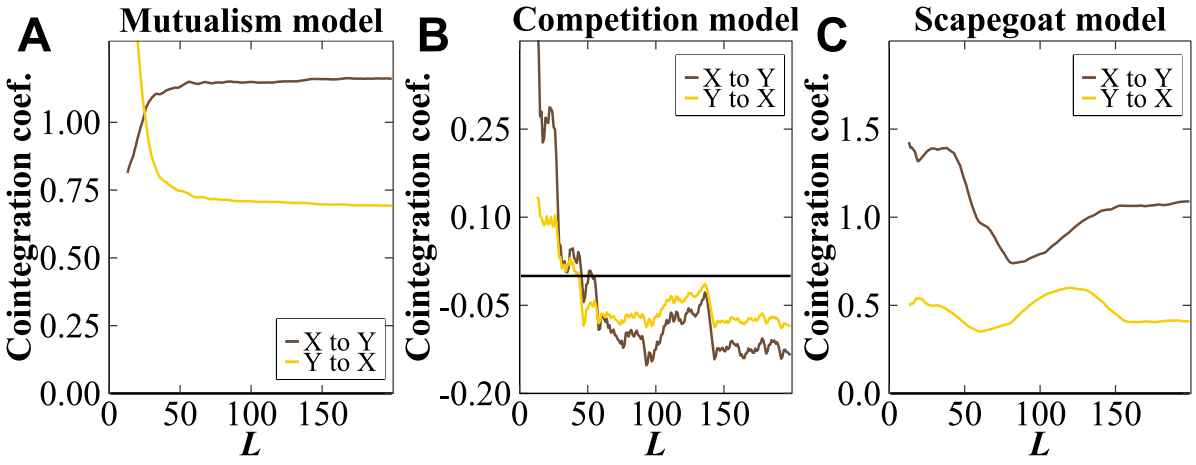


Figure 3.7: S-map interactions for the three models of positive (Eq. 3.19), negative (Eq. 3.20) and dark causality (Eq. 3.21). For the case of positive causality S-map captures clearly the influence from *X* to *Y* but is unstable from *Y* to *X*. In the case of negative causality S-map is unstable yet yields towards the negative side. Similar to the cases of correlation and cointegration, S-map is not designed to capture forms of dark causality (Fig. 3.3C).

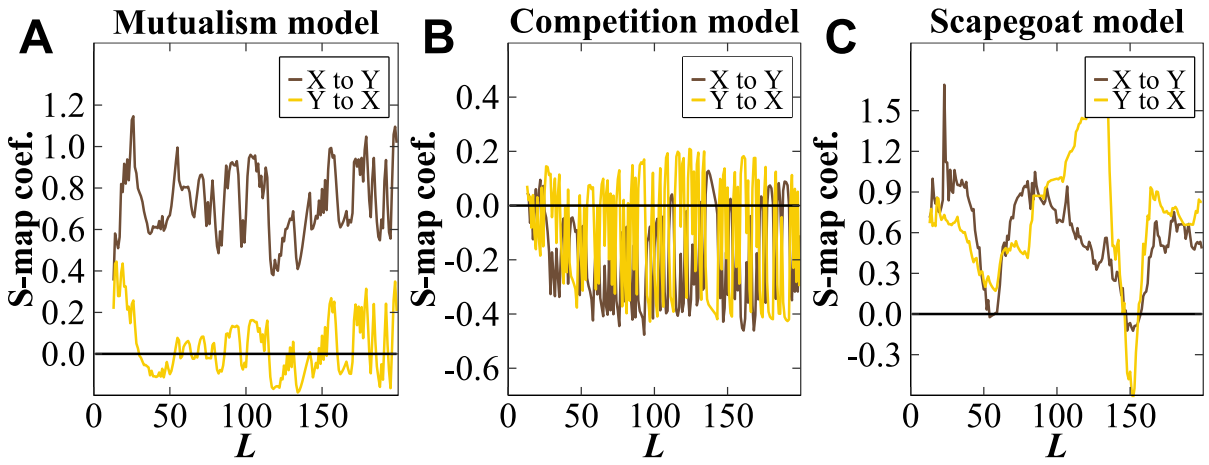
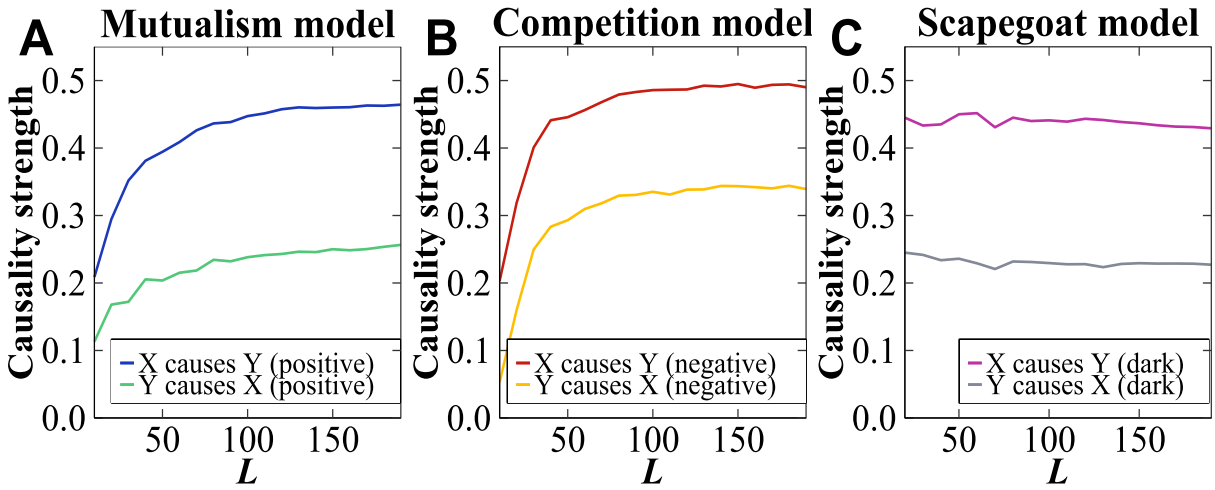


Figure 3.8: Pattern Causality for the three models of positive (Eq. 3.19), negative (Eq. 3.20) and dark causality (Eq. 3.21) demonstration.



3.7 Supplementary examples for ecology and physiology

3.7.1 Predator – prey dynamics.

To explore the nature of causality in predator – prey relations we implement PC in the standard experimental system of *Didinium* (predator) and *Paramecium* (prey) (Luckinbill 1973) (Fig. 3.9A). CCM analysis showed (Sugihara et al. 2012) that this pair of species is characterized by bidirectional causality, with causation from predator to prey being stronger than reversely.

Our results foster the top-down control from predator to prey but moreover PC displays a verbose explanation of predator-prey interactions. Fig. 3.9B reveals the distinctly negative causality from *Didinium* to *Paramecium*. This outcome is rational considering that as the predator population rises more prey gets consumed and thus prey population diminishes. Furthermore, notice that positive causality is stronger from prey to predator, which is expected if we consider that when prey population surges this supplies the predators with more to consume.

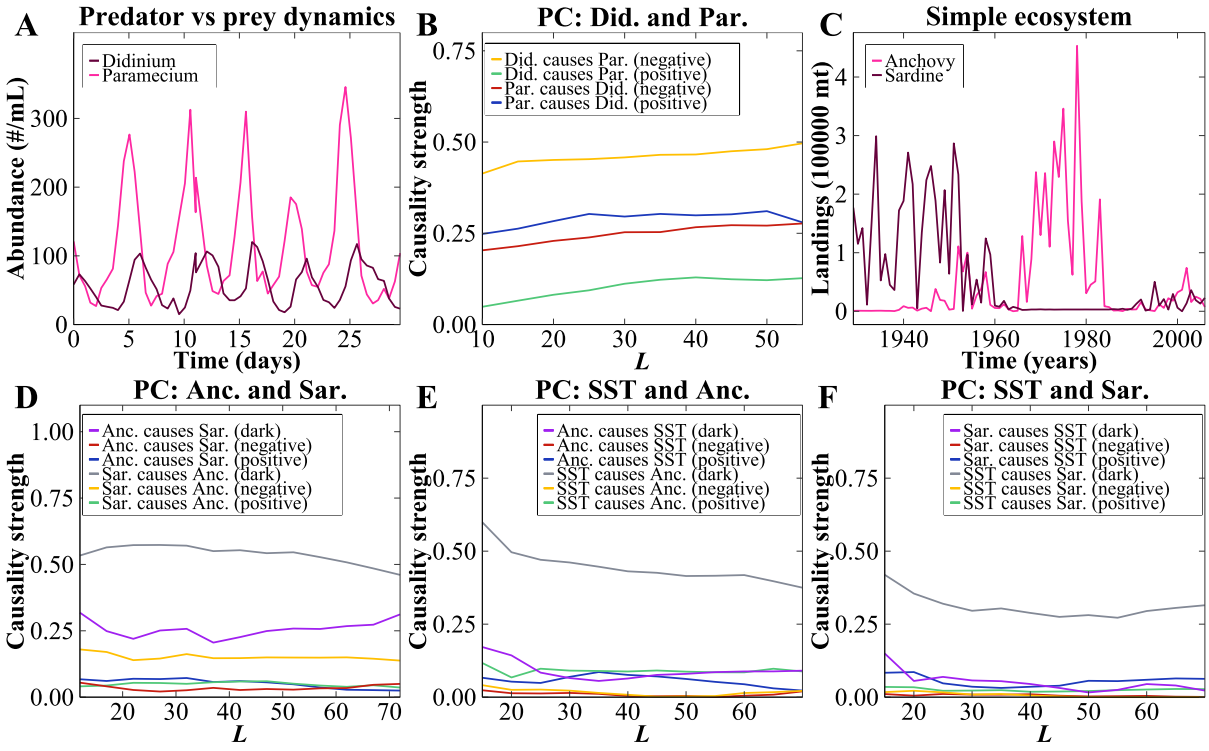
3.7.2 Hidden interactions in a simple ecosystem.

Next, we investigate an ecosystem consisting of anchovy and sardine landings (Fig. 3.9C), and sea surface temperature (SST) measured at Newport Pier and Scripps Pier, California. Such ecosystems have been a subject of conflicting analyses and mirage correlations (Jacobson and MacCall 1995; McClatchie et al. 2010). Some (Murphy and Isaacs 1964) suggest that the species are in direct competition, others (Lasker and MacCall 1983) claim that the species interaction are indirectly influenced by shared environmental forcing. CCM analysis (Sugihara et al. 2012) showed that the only causal role is that of SST influencing sardine and anchovy landings.

PC analysis exposes a complex interaction among sardines and anchovies (see Fig. 3.9D). This dark causality could possibly attest to the fact that there is controversy in the literature regarding their actual interaction (Murphy and Isaacs 1964; Lasker and MacCall 1983; Patil et al. 2001). Furthermore, we can see that causality from SST to sardines and anchovies is also of dark nature (see Fig. 3.9E, 3.9F). Therefore, we can deduce that temperature at sea level influences anchovies and sardines though in an intricate manner, neither purely beneficial nor harmful.

Overall our results suggest that apart from a dynamic (state-dependent) rule involving temperature (already advocated in Sugihara et al. 2012), acute management decisions should also be based upon closer monitoring of more subtle species interactions than obvious predation and symbiosis (e.g. resources competition).

Figure 3.9: Nature of causality in ecological data. (A) Abundance time series of Didinium and Paramecium. (B) PC exposes their mostly negative interaction. (C) Landings time series of anchovy and sardine. (D to F) Dark causality from SST towards both species as well as between the two species. Color scheme: Blue and green are used for positive causality. Red and yellow are used for negative causality. Purple and grey are used for dark causality.



3.7.3 Vital synergies in an apnea patient.

Physiological research focuses a lot on cardio – respiratory interplays to understand certain diseases (Hirsch and Bishop 1981). Sleep apnea is one of them and is severe because it causes sleep deprivation and ultimately death (Rigney et al. 1993). We investigate causal interactions among heart rate, breath rate and blood oxygen concentration (Fig. 3.10A to C) from a patient in the Beth Israel Deaconess Medical Center in Boston, Massachusetts (Rigney et al. 1993). The same dataset was studied by Schreiber (Schreiber 2000) (without considering blood

oxygen concentration) and he found information transfer to be stronger from heart to breath rate than vice versa.

By employing PC, we find that heart rate regulates breath rate in a complex way (Fig. 3.10D. Blood oxygen concentration and heart rate are involved in a mixed (positive and negative) causal loop (see Fig. 3.10E) with blood oxygen concentration yielding a stronger influence. However, the influence from blood oxygen concentration to breath rate (see Fig. 3.10F) is unilateral and of a mostly dark nature.

Our results move one step beyond Schreiber's conclusions highlighting heart rate's dominant role over breath rate (see Fig. 3.10D), and moreover unveiling the nature of their interaction. All in all, blood oxygen concentration seems to be highly influential in the cardio – respiratory dynamics and should not be neglected by medical research on apnea.

Figure 3.10: Nature of causality in physiological data of a patient with apnea. (A) Time series of heart rate (B), breath rate (C) and blood oxygen concentration. (D) Unilateral influence from heart rate towards breath rate (E, F) PC reveals the dominant causal role of blood oxygen concentration which exerts both positive and negative causality to heart and breath rate. Color scheme: Blue and green are used for positive causality. Red and yellow are used for negative causality. Purple and grey are used for dark causality.

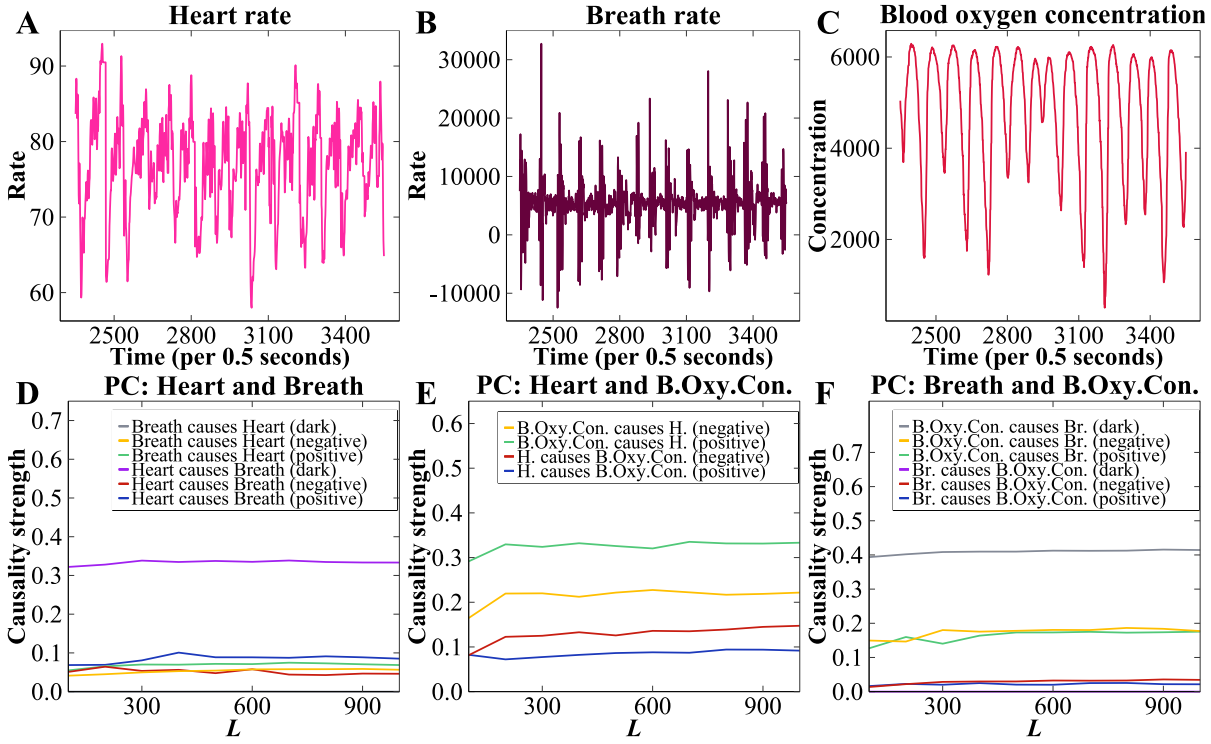


Table 3.1: PC (from X to Y) pattern to pattern matrix for E = 2.Each cell is filled with the weighted accuracy (percentage) regarding the occasions when given an average pattern in Y's neighborhood we successfully predict the average pattern in X's contemporaneous neighborhood. Thus, each cell takes values from 0 to 1. To calculate positive causality, we take the average of blue cells. For negative causality we take the average weighted accuracy inside red cells. Dark causality is the average weighted accuracy inside the purple cells.

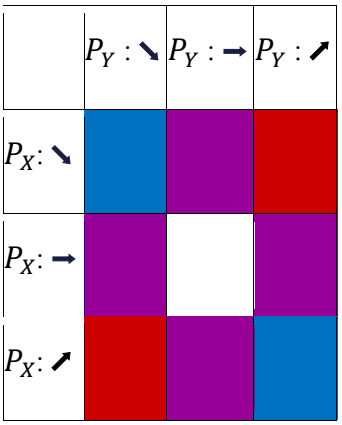
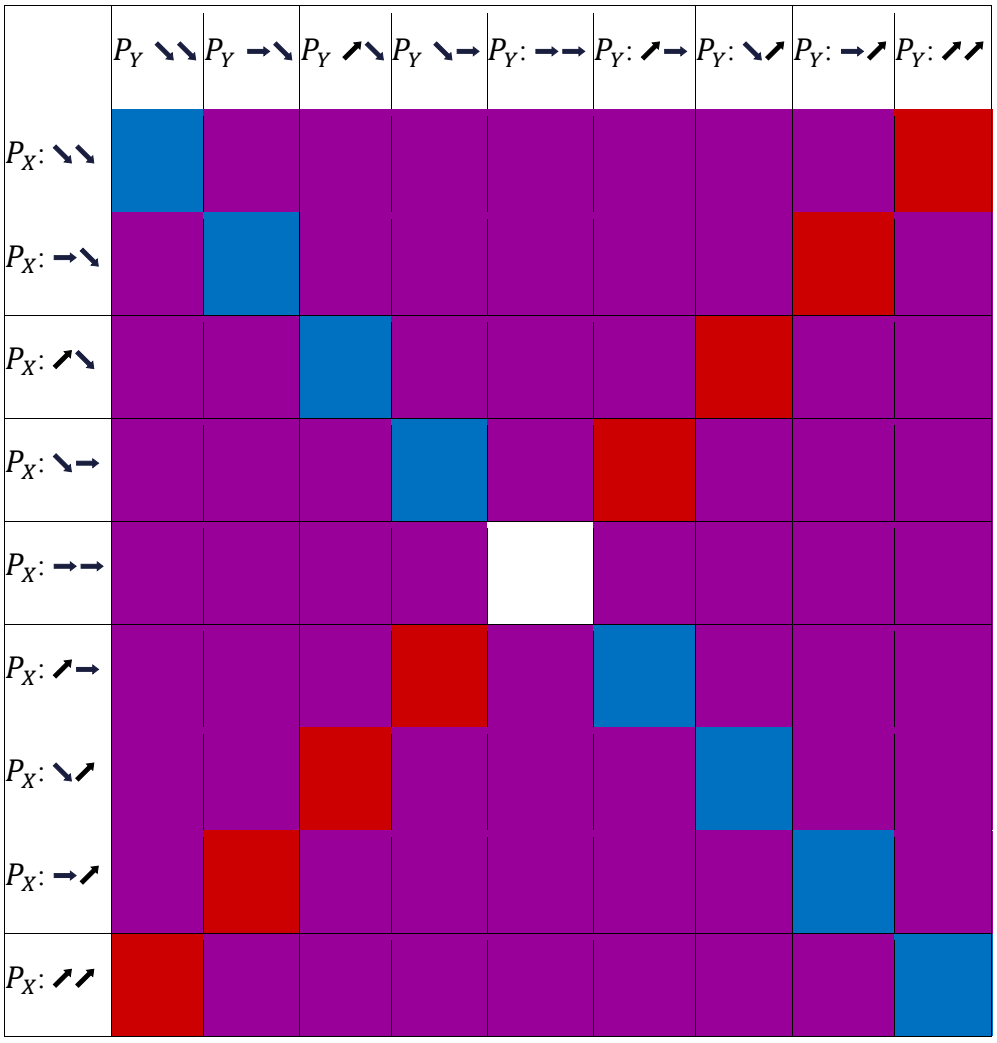


Table 3.2: PC (from X to Y) pattern to pattern matrix for E = 3.Each cell is filled with the weighted accuracy (percentage) regarding the occasions when given an average pattern in *Y*'s neighborhood we successfully predict the average pattern in *X*'s contemporaneous neighborhood. Thus, each cell takes values from 0 to 1. To calculate positive causality, we take the average of blue cells. For negative causality, we take the average weighted accuracy inside red cells. Dark causality is the average weighted accuracy inside the purple cells.



4 Unveiling Causal Interactions in Complex Systems

For centuries, philosophy illuminated the course of human endeavors. Science gave philosophy a methodological way of empirically testing theories and concepts which helped philosophers to become almost completely disentangled from superstitions and gods, seeking nature's mechanisms for the first principles of phenomena (Toulmin 1961). As an example, Thales of Miletus was able to predict the next big harvest and reserve olive presses in advance by observing the long-term impact of the weather on olive trees. Thales' predictions were accurate (Crawford and Sen 1996), as he was able to demonstrate profoundly that elaboration on the causes of things leads to a higher understanding of nature (Aristoteles and Apostle 1966). This long-standing desire to understand the first principles of phenomena provides the strongest motivation for the present study.

Natural laws govern planetary to particle motions indisputably. However, when it comes to ecosystems, brain functions and stock markets, we strive to derive first principles, causal relationships and driving factors. This lack of clear understanding is the scourge of decision and policy makers, who will eventually follow ad hoc rules or best practices (Bellman and Zadeh 1970). Unavoidably, without a clear interpretation of the systems' mechanisms and functions, fatal errors lie in wait (Russo et al. 1989). Nowadays, fortunately, the recent advances in data availability and computational power have created a fertile soil in which to develop fastidious tools for the deeper understanding of such unfathomable systems.

4.1 Experiment setups

4.1.1 Overseeing Ecosystem Interdependencies.

Ecosystems are characterized by recurring perturbations, swinging among multiple equilibria and chaotic disturbances. A small change in the native pool of species can have unpredictable impacts on the long-term balance of a given ecosystem (Chapin III et al. 2011). Environmental sentinels are concerned with species invasions and the impact of the weather on erratic regions such as desert ecosystems.

Thus, we employ our methodology in a dataset from a Chihuahuan desert scrubland site established in 1977 near Portal, Arizona, USA (Ernest et al. 2016), which contains four types of measurements: weather variables, and quantities of various rodent species, several plant species, and some ant species, a detailed list of which can be seen in Table 4.2. Our primary purpose is to retrieve, on the one hand, the causal interdependencies centered around the invader species, *Erodium cicutarium* (Valone and Balaban-Feld 2018; Valone and Balaban-Feld 2019), and on the other hand to track the traceable impact of the weather on the ecosystem.

4.1.2 Diagnosing Disorders from Brain Activity.

The brain, as a system of synaptic activity, is affected by most if not all mental disorders. For example, people suffering from alcoholism tend to exhibit adverse effects in their social life due to the neurotoxic effects on the brain, especially the frontal region. Sometimes, it even leads to persistent functional changes in brain neural circuits (Breese et al. 2011; American Psychiatric Association 2013). Principals of large-scale treatment programs can benefit from tools that are able to identify factors that differentiate afflicted subjects from control ones.

Inspired by the apparent impact of alcohol on the brain, we use a dataset made available publicly by Henri Begleiter of the Neurodynamics Laboratory of the State University of New York Health Center in Brooklyn, http://archive.ics.uci.edu/ml/datasets/EEG+Database. We use EEG measurements from 10 alcoholic and 10 control subjects. The dataset contains recordings from 64 electrodes placed on the subjects' scalps, which were sampled at 256 Hz (3.9-msec epoch) for 1 second. For our analysis, we consider each subject's exposure to a single stimulus of object pictures chosen from a curated picture set (Snodgrass and Vanderwart 1980). The electrode positions were located at standard sites (Standard Electrode Position Nomenclature according to the American Electroencephalographic Association). The data collection process is described in detail in (Zhang et al. 1995). Additionally, summary details for the electrodes corresponding to specific brain regions are provided in Table 4.3. Our purpose is to reconstruct the vital causal structure of the alcoholic brain compared to the control one.

4.1.3 Monitoring Derivatives' Systemic Risk.

Ever since the inauguration of derivative financial products, such as options and CDS, the selection and subsequent management of portfolios has become increasingly challenging. Furthermore, all market participants are intrinsically linked, and a small decision by one can

have far-reaching consequences for the market. Therefore, fund managers need to constantly investigate the ever-increasing volume of data, to optimize decision making and mitigate systemic risk.

Banks, with the incentive of hedging risk with respect to their lending operations, as well as freeing up regulatory capital, have been the prevalent actors in the CDS market. By March 1998, the global CDS market was estimated at about $300 billion, with JP Morgan alone accounting for about $50 billion of this (Tett 2006). Starting from early 2008, the global financial crisis has been quite intertwined with the role of banking CDS. Nordic and German banks have been key components of the global financial network from 2008 onwards. This motivates us to investigate further the interdependencies of banking CDS and test whether our method can identify the de facto key players during global financial crisis and post-crisis periods. We use a dataset of daily CDS spreads from the banking sector with a five-year maturity (see Table 4.4) spanning from 2007-12-14 to 2019-05-13. The time series were retrieved from Thomson Datastream.

4.2 Results

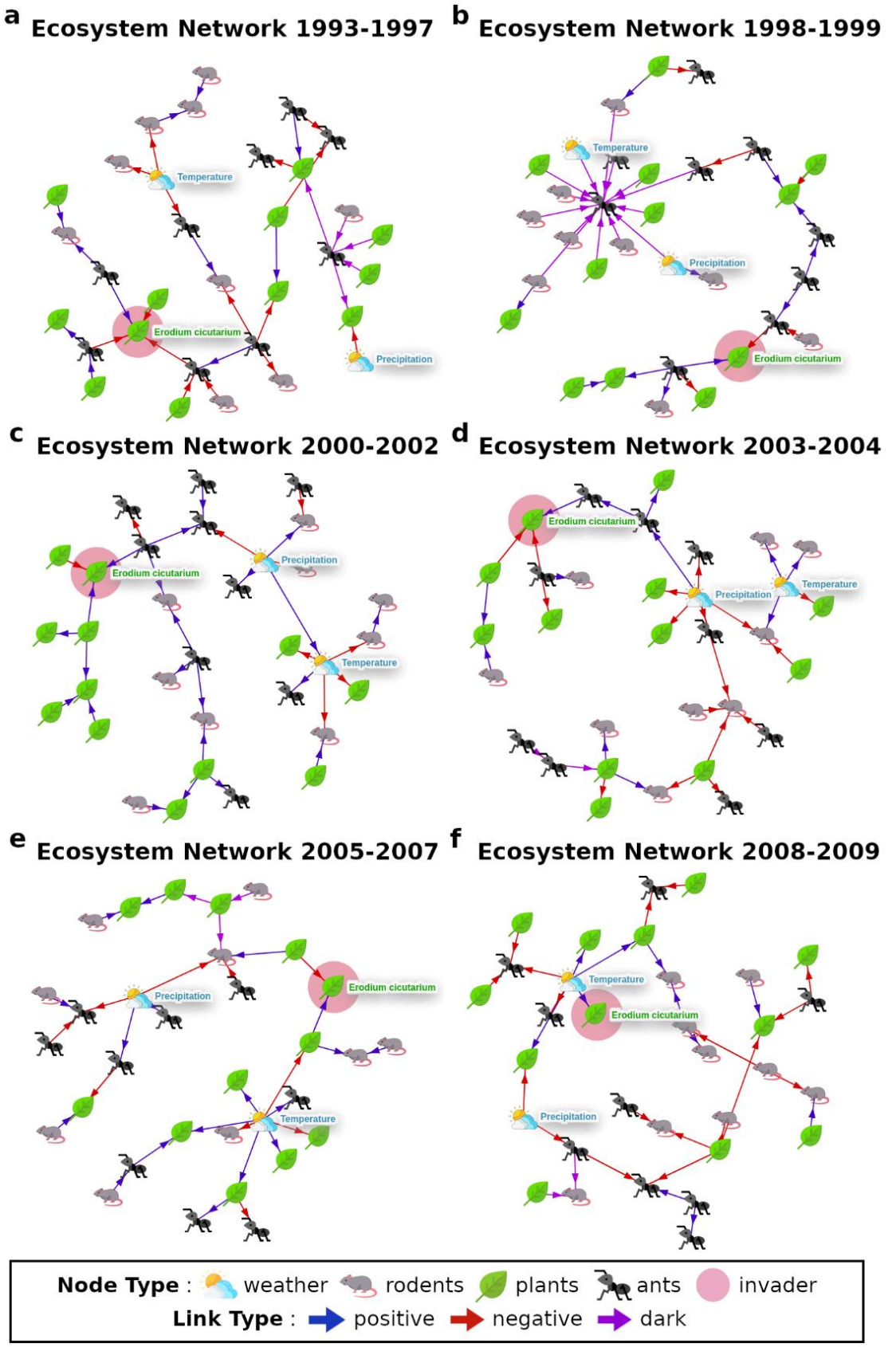
4.2.1 Tracking Invasion Dynamics and Weather Impact in a Desert Ecosystem.

During the "pre-invasion" period (Figure 4.1.a), *Erodium cicutarium* (the invader) accounted for a very small percentage of the local fauna (Valone and Balaban-Feld 2018; Valone and Balaban-Feld 2019). This information is captured with our method given that two species of ants and one species of plants are negatively related to the invader, attesting to an underlying hostility. At the "breakout" of the entrenchment (Figure 4.1.b), the invader's abundance rose to account for 25% of the fauna measured (Allington et al. 2013; Ignace and Chesson 2014), probably related to the positive influence from an ant species and the subsequent (Figure 4.1.c) positive causality from some plant and ant species. Still, though, another plant species had a negative causality on the invader, a pattern we can also see in the pre-invasion period. Later (Figure 4.1.d), despite some insisting negative influences on the invader's abundance, an ant species is found to be positively associated with the invader. Down the line (Figure 4.1.e), we find again that the invader is involved in a mixed triangle, with a plant species affecting it positively, and another ant species negatively. In the final period (Figure 4.1.f), only

temperature affects the invader, suggesting an imminent assimilation with the rest of the ecosystem. The main insight here is that sporadic positive causalities on the invader species during the post-invasion period (Figure 4.1 c-f) probably aided its successful spread in the ecosystem.

As far as the impact of the weather is concerned (Figure 4.1.a), both temperature and precipitation negatively impact two rodent species, one ant species and one plant species, attesting to the severe drought that occurred in this period (Christensen et al. 2018). Later (Figure 4.1.b), we observe the development of a dark influence regime, again involving temperature and precipitation, with an ant species. at its center. Subsequently (Figure 4.1.c), temperature and precipitation play a persistent driving role in the rest of the ecosystem, in both positive and negative ways, with precipitation later claiming more of a negative force (Figure 4.1.d) and reverting to a more balanced role thereafter (Figure 4.1.e). Ultimately (Figure 4.1.f), only temperature maintains a central role in the ecosystem, affecting plant species in a positive way. However, the fact that this period is characterized by a drought is captured by two ant species being affected by a negative causality from temperature (Pfefferbaum et al. 1997). For details on the exact species (Figures 4.4-4.9).

Figure 4.1: Aggregate causal networks for six separate periods: (**a**) Before the "aggression" of *Erodium cicutarium*, (**b-e**) Invasion periods of interest, (**f**) Post-invasion period.



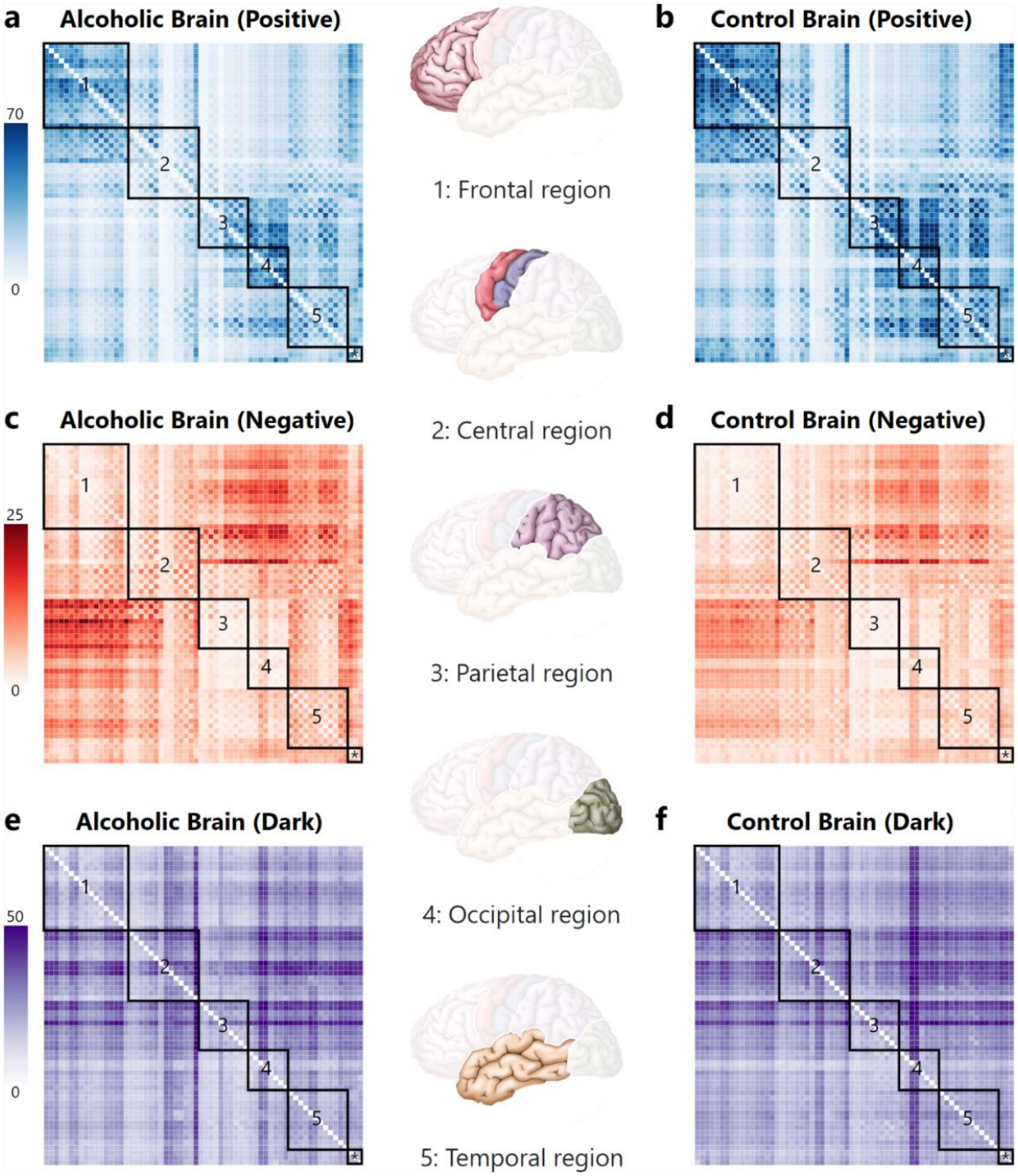
4.2.2 Revealing Distinct Features in Alcoholic Brain Networks.

In Figure 4.2, we are comparing cumulative adjacency matrices (at the end of the experiment) of the "average" alcoholic and control subjects, where the color is associated with accumulated intensity, according to Eq. (4.11) of our algorithm. Thus, darker colors denote stronger links overall. Apparently, the frontal region's positive interdependencies of the average alcoholic brain (Figure 4.2.a) are much fainter compared to the average control brain (Figure 4.2.b). This finding is related to the exhaustion of the frontal lobe due to the neurotoxic effects of alcohol (Breese et al. 2011; American Psychiatric Association 2013).

However, in terms of negative structure, it is evident that the average alcoholic brain has two specific regions in the adjacency matrix (Figure 4.2.c) with much more intense interdependencies than in the average control brain (Figure 4.2.d). These two regions translate to a negative causal regime, between frontal and parietal regions. Frontal region is responsible for the motor functions while parietal region is responsible for the perception of space as well as navigation. Our results suggest that, in the average alcoholic brain, these two regions cause opposite electrical fluctuations on each other. This is consistent with the known motor impairments as well as sensory handicaps found in an alcoholic (Pfefferbaum et al. 1997; Moselhy et al. 2001; Ratti et al. 2002).

Distinctive features are also discovered in the microstructure of dark type interactions. Most notably, in the average alcoholic brain, the voltage measurement from electrode CZ (rightmost of central region) is predicted consistently by all other electrodes' time-series (Figures 4.2.e and 4.14). This pattern is absent from the average control brain, which exhibits stronger and collective predictive capacity on electrodes PO7 and PO8 (Figures 4.2.f and 4.15), which are associated with properties related to visual memory (occipital region). Interestingly enough, the occipital region is involved in the processing of the object pictures, the region of interest in this experiment. Our analysis suggests a higher influence of occipital region from all brain regions in the control brain, a fact already reflected in the brain research literature (Zhang et al. 1995; Begleiter et al. 1993; Hertz et al. 1994). For details on the exact electrodes (Figures 4.10-4.15).

Figure 4.2: Cumulative adjacency matrices for the average positive/negative/dark network structures of alcoholic (a, c, e) and control subjects (b, d, f), for the whole experiment duration. Darker color denotes higher accumulated link strength.



4.2.3 Detecting Persistent Causal Relationships and Influential Assets in the CDS Market.

The most straightforward way to rank CDSs' contribution to systemic portfolio risk is via influence exerted and influence received. Effectively, we can become aware of which are the CDS that influence others, while at the same time receiving less influence.

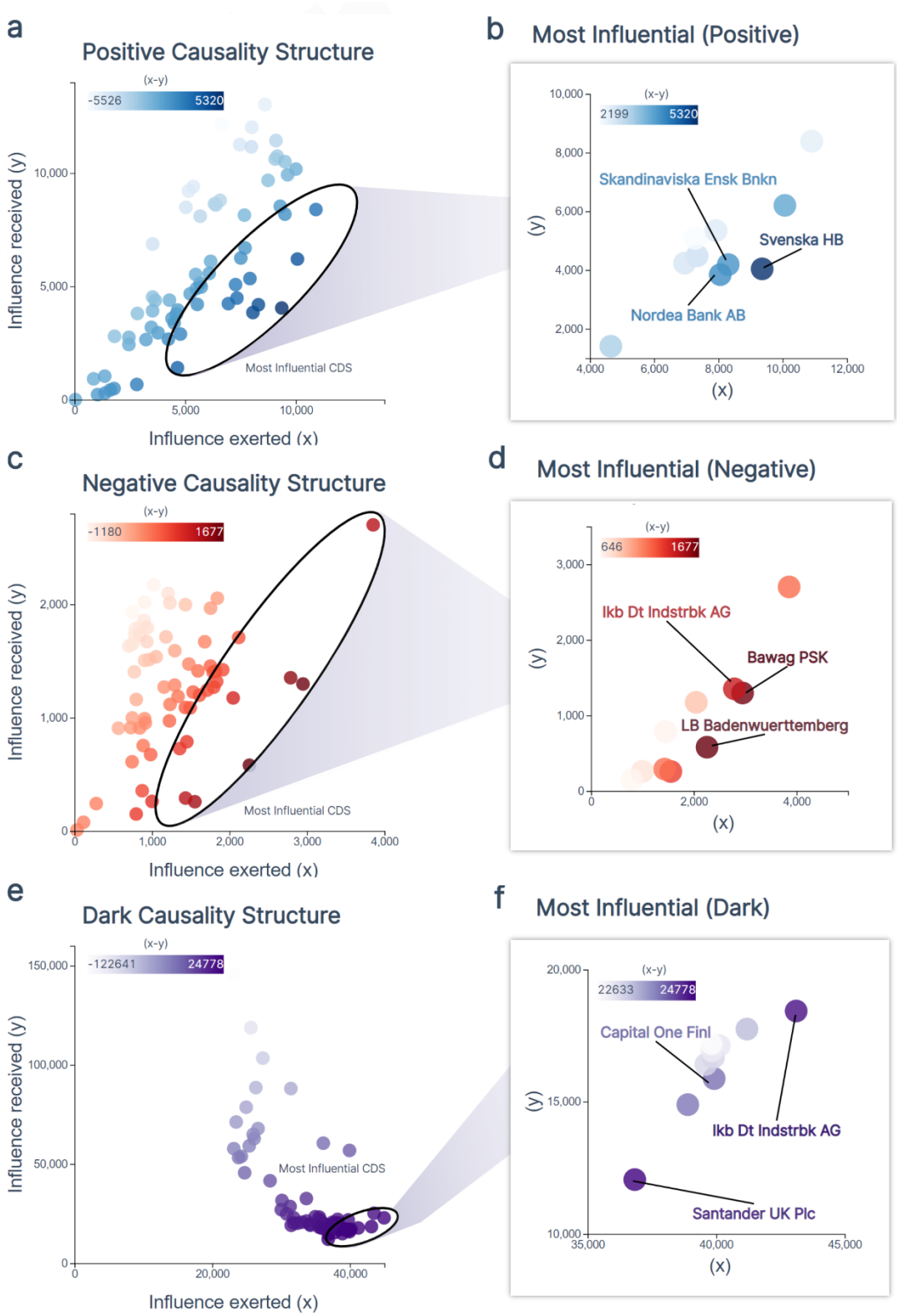
We observe that, in terms of positive interdependencies (Figure 4.3a), the layout of the CDS causality structure seems to be arranged in a homogeneous manner, suggesting that, when considering both exerted and received influence, the majority of CDS seem to exhibit a balance between the two. Notably, the most influential CDS are Svenska Handelsbanken, Nordea Bank AB and Skandinaviska Ensk Banken (Figure 4.3b, see also Table 4.5 for the top 10). This result suggests that the specific Nordic banks' CDS had the highest same-direction predictive capacity for the rest of the CDS in our dataset. This result might associate with the fact that the Nordic banks were experiencing higher loan-to-deposit ratios than all other banks, leaving them quite exposed to systemic risk, an event which could turn the market's eyes on their credibility, thus making their CDS spreads quite the driving market force (Babihuga and Spaltro 2014). Moreover, Svenska Handelsbanken has been a center of attention in terms of its innovative banking model (Kroner 2011).

A similar structure is evident with regards to negative interdependencies (Figure 4.3c), though a bit more dispersed, implying a sharper difference between influence exerted and received. The most influential here are Landesbank Badenwuerttemberg, Bawag PSK and Ikb Deutschet Industriebank AG (Figure 4.3d, see also Table 4.5 for the top 10). Notably, these are German banks, and, in the period under study, they were found to hold huge amounts of sovereign bonds in their balance sheets (Buch et al. 2016), effectively making them the biggest players in the sovereign derivatives market.

Ultimately, contemplating Figure 4.3e, we deduce that the dark causality structure of the dark interdependencies is different compared to the ones observed in the positive and negative interdependencies. CDS receiving much influence, exert much less, while CDS exerting much influence, receive less, compared to the previous two cases (positive and negative). In this case, the most influential CDS are those of Santander UK PLC, Ikb Deutschet Industriebank AG and Capital One Financial (Figure 4.3f, see also Table 4.5 for the top 10). At first sight, these banks seem unrelated. However, they were found to be at the very center of the 2007-2008 crisis

(Erkens et al. 2012; Grammatikos and Vermeulen 2012). For details on the exact CDS interdependencies (Figures 4.16-4.18 and Table 4.5).

Figure 4.3: Bubble Plot of CDS in terms of positive exerted and received influence (out and in strength centrality respectively).Size of points is analogous to out-strength minus in-strength. Points closer to orange receive relatively more influence while points closer to blue exert relatively more influence.



4.3 Concluding remarks

In this work, we introduce a novel framework for the detection of latent and elusive structures in causal networks. Our method is based on short-term predictions drawn from information embedded in a reconstructed state space. The prudent algorithmic design reveals time-series causalities in three echelons, i.e., positive (same-direction), negative (opposite-direction) and dark (mixed-direction) predictive relationships. This targeted partition allows the unique identification of persistent causal structures and dominant influences which would otherwise be lost in the noise of disparate causalities (if we did not discern the three types of interactions). Applying this method to a set of time-series measurements from a given complex system allows us to perceive deeply rooted causalities for each of the three types separately. We demonstrate our method's power to discern the most fundamental components, i.e. "the backbone" that drives a system's evolution in three different disciplines.

As a first challenge, we tested our method on a desert ecosystem with imperfect measurements of weather conditions, as well as abundances of rodents, plants and ants. From observation (Valone and Balaban-Feld 2018; Valone and Balaban-Feld 2019), it was known that this ecosystem experienced an exotic species invasion as well as two periods of severe drought. Our method was able to quantitatively capture the invasion's dynamics as well as some extra information regarding possible "inside assistance" for the invader species. Moreover, the central role of the weather, both during the droughts and in the other periods, was effectively in tandem with empirical findings (Christensen et al. 2018). Next, we tested our method on a setting from neurology. Well-established literature (Pfefferbaum et al. 1997; Moselhy et al. 2001; Ratti et al. 2002) had noted alcoholism's impact on the frontal lobe. Through our method we found much fainter positive interdependencies in the frontal region of the average alcoholic compared to the average control brain. Furthermore, under the dark causality spectrum, we were able to identify the control brain's higher activity in the occipital region (visual cortex). Finally, being aware of specific banks' highlighted roles during the last decade (Babihuga and Spaltro 2014; Kroner 2011), we wanted to test our method's capacity to reconstruct the CDS causal network, while capturing the most impactful components. Indeed, our method was able to identify the high impact of Nordic and German banks on the rest of the banks' CDS, as well as banks whose role was very central during the 2007-2008 financial crisis. In each case, we are able to reveal the most important factors driving the rest of the system under scrutiny.

Finally, the proposed method is able to capture a range of causal links in a variety of complex systems. However, from these perspectives, we would like to see the application of the suggested novel methodology beyond the presented examples, and its reach extended to a much broader class of topics.

4.4 Methods

Before the theoretical methodology is developed to reveal causal networks, it is necessary to introduce the following notation:

4.4.1 A framework of causality assessment.

The predictive capability of this novel approach is assessed by establishing a causal relationship between time series. While, in (Stavroglou et al. 2019) the influence from X to Y is merely quantified by comparing patterns from contemporaneous neighborhoods of M_X and M_Y, here we investigate this relationship further using patterns from M_X's current neighborhood to *predict* M_Y's future patterns (h steps ahead of time t). In other words, the strong predictive power of our treatment is deployed by the algorithm formulated below. However, to demonstrate the worth of our contribution, the mathematical formalities (i.e., lemmas, theorems and their proofs) are delineated in section 4.10. In particular, in what follows, according to Lemma 4.1, M_X is said to strongly influence M_Y in an absolute way if all values of M_Y are affected by M_X, which will be tested each time we accurately predict a future pattern of M_Y, i.e. when Eq. (4.4) equals Eq. (4.9). Further, the strength of the influence is calculated by the intensity ratio, see Eq. (4.10), and we expect Lemma 4.3 to hold, which states that some (and not all) of M_Y's values are affected by M_X. Lemmas 4.2 and 4.4 suggest that, if M_X influences M_Y, then subsequently X influences Y, effectively allowing conclusions from attractor analysis to be interpreted for raw time series as well. Finally, Theorems 4.1, 4.2 and 4.3 respectively separate the nature of influence into positive, negative and dark, and they are included at the end of our method when we use the PC matrix (see Tables 4.6 and 4.7) to support the visualization of our treatment.

4.4.2 Shadow Attractors Reconstruction.

We create the shadow attractors, M_X and M_Y, for X and Y, respectively, by finding the optimal pair (E, τ). In particular, we initially compare the predicting accuracy for a whole range of reasonable embedding values of E and τ, and then we calculate the distance matrices, D_X and D_Y (e.g. either using the L_1 norm if we want to treat all distances equally or L_2 if we want to penalize bigger distances), among all vectors in M_X and M_Y.

$$X = \{X(1), \dots, X(L)\} \Longrightarrow M_X = \begin{pmatrix} \underline{x}(1) =< X(1), X(1+\tau), \dots, X(1+(E-1)\tau) > \\ \underline{x}(2) =< X(2), X(2+\tau), \dots, X(2+(E-1)\tau) > \\ \vdots \\ \underline{x}(L-(E-1)\tau) =< X(L-(E-1)\tau), X(L-(E-2)\tau), \dots, X(L) > \end{pmatrix}$$

and

$$D_X = \begin{pmatrix} d\left(\underline{x}(1), \underline{x}(1)\right) & \cdots & d\left(\underline{x}(1), \underline{x}(L-(E-1)\tau)\right) \\ \vdots & \ddots & \vdots \\ d\left(\underline{x}(L-(E-1)\tau), \underline{x}(1)\right) & \cdots & d\left(\underline{x}(L-(E-1)\tau), \underline{x}(L-(E-1)\tau)\right) \end{pmatrix}. \quad (4.1)$$

We derive M_Y and D_Y similarly.

Once the shadow attractors are derived, we obtain access to the reconstructed topology of the complex system. In the next step, we parse the local areas in the attractors and extract useful information for the prediction and the causality inference.

4.4.3 The Nearest Neighbors and their Future Projections.

For each point $\underline{x}(t)$ in M_X, we find its $E+1$ nearest neighbors $NN_{\underline{x}(t)}$, which is the minimum number of points needed for a bounded simplex in an E-dimensional space. From these $E+1$ nearest neighbors, we need to keep the time indices, find the corresponding points on M_Y, and project them ahead by *h* steps to determine the future states:

a. The projected time indices $t_{\underline{x}_1}, t_{\underline{x}_2}, \dots, t_{\underline{x}_{E+1}}$:

$$NN_{\underline{x}(t)} = argmin_{(E+1)} \left\{ d\left(\underline{x}(t), \underline{x}(1)\right), \dots, d\left(\underline{x}(t), \underline{x}(t-(E-1)*\tau-h)\right) \right\} =$$

$$\{NN_{\underline{x}(t_1)}, NN_{\underline{x}(t_2)}, \dots, NN_{\underline{x}(t_{E+1})}\} \Longrightarrow t_1, t_2, \dots, t_{E+1} \xrightarrow{projecting\ h\ steps\ ahead} t_1 + h, t_2 + h, \dots, t_{E+1} + h = t_{\underline{x}_1}, t_{\underline{x}_2}, \dots, t_{\underline{x}_{E+1}} \quad (4.2)$$

b. The distance of the projected neighbors from $\underline{y}(t)$:

$$d_{\underline{y}_1} = d(\underline{y}(t), \underline{y}(t_{\underline{x}_1})), d_{\underline{y}_2} = d(\underline{y}(t), \underline{y}(t_{\underline{x}_2})), \ldots, d_{\underline{y}_{E+1}} = d(\underline{y}(t), \underline{y}(t_{\underline{x}_{E+1}})), \quad (4.3)$$

In order to avoid any data snooping, the following must hold for all the projections of the nearest neighbors: $t_n < t$, where $t_n \in \{t_{\underline{x}_1}, t_{\underline{x}_2}, \ldots, t_{\underline{x}_{E+1}}\}$. In this step, we extract the projected time indices of $\underline{x}(t)$'s neighbors' projections and use them to calculate the distances of their co-temporals $\underline{y}(t_{\underline{x}_n})$, where $t_{\underline{x}_n} := t_{\underline{x}_1}, t_{\underline{x}_2}, \ldots, t_{\underline{x}_{E+1}}$.

4.4.4 The Affected Variable's Predicted Pattern *h* Steps Ahead.

We use the relevant information from Eqs. (4.2) and (4.3) to estimate the predicted pattern $\hat{P}_{\underline{y}(t+h)}$ of $\underline{y}(t+h)$:

$$\hat{P}_{\underline{y}(t+h)} = signature(\hat{S}_{\underline{y}(t+h)}), \quad (4.4)$$

where

$$\hat{S}_{\underline{y}(t+h)} = \sum_{t_n=t_{\underline{x}_1}}^{t_{\underline{x}_{E+1}}} w_{t_n}^{\underline{x}} s_{t_n}^{\underline{y}}, \quad (4.5)$$

$$w_{t_n}^{\underline{x}} = \frac{e^{d_{t_n}}}{\sum_{t_n} e^{d_{t_n}}}. \quad (4.6)$$

Here, the d_{t_n} represent the distances from Eq. (4.3), and

$$s_{t_n}^{\underline{y}} = \left(\frac{\mathrm{Y}(t_{\underline{x}_2}) - \mathrm{Y}(t_{\underline{x}_1})}{\mathrm{Y}(t_{\underline{x}_1})}, \ldots, \frac{\mathrm{Y}(t_{\underline{x}_{E+1}}) - \mathrm{Y}(t_{\underline{x}_E})}{\mathrm{Y}(t_{\underline{x}_E})}\right). \quad (4.7)$$

Remark: $t_{\underline{x}_1}, t_{\underline{x}_2}, \ldots, t_{\underline{x}_{E+1}}$ correspond to the ones calculated in Eq. (4.2).

Here, we are using information from M_X in order to predict M_Y's future pattern $\underline{y}(t+h)$.

4.4.5 The Driver Variable's Pattern.

Then, we keep the current pattern of $\underline{x}(t)$, which is $P_{\underline{x}(t)}$:

$$P_{\underline{x}(t)} = signature(S_{\underline{x}(t)}), \quad (4.8)$$

where the signature is the way of extracting patterns from vectors.

By holding the current signature of $\underline{x}(t)$, we are able to assess both the intensity and the type of the causality from X to Y.

4.4.6 The Affected Variable's Real Pattern (Backtesting Process).

Then, we keep the real pattern of $\underline{y}(t+h)$, which is $P_{\underline{y}(t+h)}$:

$$P_{\underline{y}(t+h)} = signature(S_{\underline{y}(t+h)}). \tag{4.9}$$

Here, we extract the real signature of $\underline{y}(t+h)$ and we are able to test our hypothesis for causality. In order for that to be true, the predicted pattern from Eq. (4.4) must be the same as the real pattern from Eq. (4.9). This process is in accordance with Lemmas 4.1 and 4.3.

4.4.7 The Nature and Intensity of Influence at Every Time Step *t*.

We repeat this procedure, see Eqs. (4.2) - (4.9), for every point of the shadow manifold M_X and fill in the PC matrix (see Tables 4.6 and 4.7) for every time step t whose influence is valid as described above. Otherwise, the PC matrix for the current t is left empty. We fill in the PC matrix, when the prediction is valid, by calculating the norms of the signatures, which are the representations of the pattern's strength, and divide the cause's norm $\left\|S_{\underline{x}(t)}\right\|$ by the effect's norm $\left\|S_{\underline{y}(t+h)}\right\|$:

$$PC[P_X, P_Y, t] = \frac{\left\|S_{\underline{y}(t+h)}\right\|}{\left\|S_{\underline{x}(t)}\right\|}. \tag{4.10}$$

For a normalized output, we can instead fill in the PC matrix by filtering first with the error function

$$PC[P_X, P_Y, t] = erf\left(\frac{\left\|S_{y(t+h)}\right\|}{\left\|S_{\underline{x}(t)}\right\|}\right), \tag{4.11}$$

where erf (·) stands for the error squashing function.

4.4.8 The Overall (for All *t*) Nature and Intensity of Causality.

At this point, the produced results contain three time series, one for each type of influence (positive, negative and dark), labelled $P(t), N(t)$ and $D(t)$ respectively, indicating at each time step the intensity of the influence (from 0 to 1). Notice that, for a given t, only one of the three can be different than zero, meaning that we cannot have more than one type of influence at the same time.

4.4.9 Causal Network Analytics.

Doing research in the era of big data involves the analysis of interdependencies among many time-series variables. Thus, instead of just X and Y, we have N variables, i.e. $X_1, \dots, X_N$. The variables are heretofore referred to as "*nodes*" of a network. Hence, the maximum number of causal interactions to be put under scrutiny is $N(N-1)$, not accounting for loops. Now, we can have a total of $N(N-1)$ resulting time series of each type (referring to $P(t), N(t)$ and $D(t)$), effectively creating three dynamic causal networks, one for each aspect (positive, negative and dark), or symbolically:

- $P_k^l(t)$, referring to the intensity of positive influence at time t, from node k to node l,
- $N_k^l(t)$, referring to the intensity of negative influence at time t, from node k to node l,
- $D_k^l(t)$, referring to the intensity of dark influence at time t, from node k to node l.

Ultimately, $P_k^l(t), N_k^l(t), D_k^l(t)\ \forall k, l$ are the positive, negative and dark aspects, respectively, of the causal network at time t, and can be seen as three concurrent networks, of the same nodes but with mutually exclusive links (no link can exist at the same time for more than one of the three aspects). Optionally, we can filter the network to keep only the strongest relationships by using algorithms such as the Minimum/Maximum Spanning Tree (Hu 1961; Gower and Ross 1969) or the Planar Maximally Filtered Graph (Tumminello et al. 2005).

Strength Centrality. This metric refers to the aggregation of the weights of the links from and to the node (Barrat et al. 2004). Out-strength denotes the weighted influence exerted directly on other nodes and in-strength denotes the weighted influence received directly from other nodes. Weights, here, are calculated from Eq. (4.11).

Link Persistence. This measures the overall weight of a given link from node X to node Y by aggregating cumulatively across time to rank time-series interdependencies on strength and persistence (Stavroglou et al. 2017).

Complexity. The proposed method is computationally efficient for long time series (large L). The only parameters that impact our method are the time series length L and its embedding dimension E. The higher is L and/or E, the longer it will take for the distance matrices D_X and D_Y to be calculated. To extract the candidate neighbors of a point $\underline{x}(t)$, we only need the $D_X[t, 1:(t-1)]$ part of D_X (same for D_Y). Computing D_X and D_Y costs L^2E for each, and the iteration part of the main algorithm is of order $O(L)$. The total cost of our algorithm is of order $O(L^2E + L)$, with the main bulk of the calculations being that of the initial distance matrices.

Method's Validation using Simulation. Our novel method has been validated using 100,000 simulations and different lengths of chains for the three types of interactions, positive, negative and dark. Analysis and discussion of this simulation-based validation is provided below. Particularly for short chain lengths, the results derived are rather impressive.

4.5 Extensive Algorithm Plug-ins

4.5.1 Node-level statistics

- *degree* is the number of outgoing and incoming links in a given node (Hakimi 1962). Out-degree denotes the influence exerted directly on other nodes and in-degree denotes the influence received directly from other nodes.
- *closeness* measures the reciprocal of how many steps are required to access every other node from a given node (Bavelas 1950). High out-closeness means that a node might spill its influence over onto further nodes (apart from its direct links). Similarly, high in-closeness suggests that a node might receive indirect influences from nodes beyond its immediate proximity.
- *betweenness* is (roughly) defined as the number of geodesics (shortest paths) going through a node (Freeman 1977). High betweenness in our framework suggests that the nodes in question might enable the propagation of indirect influences across the whole network.
- *eigenvector* measures the influence a node has on a network. If a node is pointed to by many nodes (which have high eigenvector centrality), then that node will also have high

eigenvector centrality (Newman 2016). Practically speaking, a high-eigenvector node in an interdependencies network indicates a time series which is at the top of the influential hierarchy.

- *strength* is the aggregation of the weights of the links from and to the node (weighted degree) (Barrat et al. 2004). Out-strength denotes the weighted influence exerted directly on other nodes and in-strength denotes the weighted influence received directly from other nodes. Weights, here, are obtained from Eq. (4.11).
- *node diversity* measures the diversity of a given node according to some node property (e.g. degree) (Eagle et al. 2010). In our framework, we can use this statistic to identify time series with strange or uncommon interdependencies structures.

4.5.2 Link statistics

link persistence measures the overall weight of a given link from node X to node Y by aggregating cumulatively across time in order to rank time series interdependencies by strength and persistence (Stavroglou et al. 2017).

4.5.3 Neighborhood-level statistics

- *rich club coefficient* measures the extent to which nodes with many links also connect to each other (Zhou and Mondragon 2004). In our framework, we can extract two types of rich clubs: (a) one considering only outgoing links, allowing us to locate which time series constitute the "main driving core" of the system, and (b) another considering only incoming links, allowing us to locate which time series constitute the "main affected core" of the system. We refer to these rich clubs collectively as "hyper-active cores".
- *k-core decomposition*: the k-core of a network is a maximal subgraph in which each node has at least degree k. The coreness of a node is k if it belongs to the k-core but not to the $(k+1)$-core (Seidman 1983). This statistic can give us the detailed hierarchical structures of influence exertion (considering out-degree) and influence absorption (considering in-degree).
- *community detection* identifies sets of nodes such that each set of nodes is densely connected internally (Porter et al. 2009). Communities in interdependencies networks suggest groups of time series which are almost independent from the time series outside their community.

4.5.4 Network-wide statistics

- *aggregate intensity* is simply the summation of all the link's weights, as calculated from Eq. (4.11). It denotes the overall influence at play in a given network and can be used to compare different networks in a straightforward manner.
- *centralization* measures how central the most central node is in relation to how central all the other nodes are (Freeman 1978). Centrality is measured according to the aforementioned node-level statistics. A highly centralized (in terms of out-degree for example) causal network suggests the existence of a time series acting as a singularity of influence over the whole set of time series.
- *components* measures the number of weakly connected components (Hopcroft and Tarjan 1973). The higher is the number of components, the less global the influence exerted can be considered, meaning that the time series under study cluster in smaller groups.
- *density* measures how densely connected the network is (Coleman and More 1983). High density of a causal network suggests intense bursts of influence among all its time series.
- *articulation* measures the percentage of articulation nodes (Barnes and Harary 1983). Such nodes (in our case, time series), if removed, disconnect the network, meaning they probably play a connector role among densely interdependent time series. Articulation nodes are, in brief, bridges of influence.
- *average path length* measures the average number of nodes needed for each node to reach every other node (West 1996). The lower it is, the more direct are the interdependencies in a system.
- *clustering coefficient* measures the degree to which nodes in a graph tend to cluster together (Luce and Perry 1949). High clustering suggests many triangles of interdependencies.
- *entropy*: the extent to which the frequency distribution of a node property (e.g. degree centrality) is uniform (Shannon 1948). Picking a node-level statistic and using entropy over its distribution across all nodes allows us to understand how diverse (low entropy) or uniform (high entropy) is the specific influence characteristic in our system.
- *modularity (walktrap)* measures the strength of division of a network into communities. High modularity suggests dense connections between the nodes within communities but sparse connections between nodes in different communities. The walktrap algorithm works well with directed networks which are the natural product of pattern causality (Pons and Latapy 2006).

- *assortativity coefficient* takes positive values if similar nodes (based on a node property) tend to connect to each other, and negative values otherwise (Newman 2002). If a causal network has a high assortativity coefficient, then it means that time series of similar interdependencies cluster together.
- *scale-free property* is the property of a network whose degree distribution follows a power law, at least asymptotically (Price 1965). In our framework, a network being scale-free in terms of out-degree would suggest the existence of few time series with highly influential connectivity and many with weak influential connectivity. Analogous conclusions would stem from a network being scale-free in terms of in-degree.
- *small world property* characterizes a network in which most nodes are not neighbors of one another, but the neighbors of any given node are likely to be neighbors of each other and most nodes can be reached from any other node in a small number of hops or steps (Watts and Strogatz 1998).

4.6 Big "O" Complexity

Our method is computationally efficient for long time series (large L). Note that our algorithm receives as inputs E and τ and the pair of time series under consideration. τ does not affect the complexity of the algorithm. Therefore, the only parameters that impact our method are the time series length L and its embedding dimension E. The higher are L and/or E, the longer it will take for the distance matrices D_X and D_Y to be calculated. A simple trick to avoid the mass calculations of D_X and D_Y at each iteration (for every time step) is to compute them at the beginning, before any iteration, once and for all. This way, to extract the candidate neighbors of a point $\underline{x}(t)$, we only need the $D_X[t, 1{:}\,(t-1)]$ part of D_X (same for D_Y). Computing D_X and D_Y costs L^2E for each and the iteration part of the main algorithm is of order $O(L)$.

The total cost of our algorithm is of order $O(L^2E + L)$ with the main bulk of the calculations being that of the initial distance matrices.

Below, we lay out the analytical complexity for every step of our algorithm:

Step 1: Reconstructing the shadow attractors M_X and M_Y is of order $O(L)$ and calculating D_X and D_Y is of order $O(L^2E)$.

Steps 2 to 6 are performed for every t in the time series. Therefore, the loop including them is of order $O(L)$.

Step 2: Parsing every row (which corresponds to finding the nearest neighbors for each point $\underline{x}(t)$) in D_X is of order $O(E+1)$.

Step 3: The extraction of $\hat{P}_{\underline{y}(t+h)}$ is trivially of order $O(E+1)$.

Step 4: Retrieving $P_{\underline{x}(t)}$ is of order $O(E)$.

Step 5: Retrieving the real $P_{\underline{y}(t+h)}$ pattern is of order $O(E)$.

Step 6: Filling the PC pattern into the pattern matrix is a procedure of order $O(1)$.

4.7 Synthetic Interdependencies Validation

The purpose of this application is to validate the capacity of our method to discern positive, negative and dark interdependencies. We use the following toy model to generate a synthetic influence from a time series X to a time series Y:

$$X(t) = X(t-1) + S_{\underline{x}(t)}, \qquad S_{\underline{x}(t)} \in \mathbb{R}^{E-1}$$

$$Y(t) = Y(t-1) + \alpha(X(t), S_{\underline{x}(t)}, \kappa),$$

where

$$\alpha\big(X(t), S_{\underline{x}(t)}, \kappa\big) = \begin{cases} S_{\underline{y}(t)} : P_{\underline{y}(t)} \parallel P_{\underline{x}(t)}, & \kappa = "positive" \\ S_{\underline{y}(t)} : P_{\underline{y}(t)} \perp P_{\underline{x}(t)}, & \kappa = "negative" \\ S_{\underline{y}(t)} : P_{\underline{y}(t)} \nparallel P_{\underline{x}(t)}, & \kappa = "dark" \end{cases}$$

We have $S_{\underline{x}(t)}$ following a uniform distribution and $S_{\underline{y}(t)}$ taking values according to the function $\alpha\big(X(t), S_{\underline{x}(t)}, \kappa\big)$ described above. We run 100,000 simulations separately for each of the three types of influence (positive, negative, dark) we want to synthesize. For this application, we use $E = 2, \tau = 1$ and $h = 1$.

4.7.1 Synthetic influences generation

To simulate a positive influence, for each time step *t* we use a uniform *random number generator* (RNG) with the following settings: 90% chance that κ = "positive", 5% chance that κ = "negative" and 5% chance that κ = "dark". The reason we do this is to allow our model to produce mixed influence types in order to make it more realistic and test our method in a more robust setting.

Similarly, we simulate a negative influence by giving the RNG the following settings: 5% chance that κ = "positive", 90% chance that κ = "negative" and 5% chance that κ = "dark". A dark influence is generated as follows: 5% chance κ = "positive", 5% chance κ = "negative" and 90% chance κ = "dark".

4.7.2 Recording chains of consecutive influences of a single type of influence

After producing 100,000 simulations for each type of influence, we record consecutive incidents or chains (up to a length of 15) of the same type we want to test (e.g. for positive influence simulations we consider only positive instances of influence etc.), and we record in columns A, C and E of Table 4.2 the percentage our toy model generated (for each chain length). For example, in the positive simulations, positive chains of length 2 were recorded in 85.742% (Col. A, Chain 2 row) of cases in the 100,000 simulations, meaning that the remaining 14.258%are instances of negative and dark influence.

4.7.3 Measuring the accuracy of our algorithm in detecting the target type of influence

Once we have determined the percentages of chains of interest, we run our method in order to gauge the extent to which it can detect the actual type of influence for chains of lengths up to 15. As we can observe from Table 4.2, for the positive influence test (Col. B) our algorithm reaches 93.5685% accuracy in terms of single positive instances (chain length = 1) and dwindles smoothly to a floor of 82.9421% accuracy for positive chains of length 15. For the case of negative simulations, our method performs equally well (Col. D), starting from 93.4461% for single negative instances and ending up at an 84.7194% accuracy level for negative chains of length 15. Lastly, in the case of dark simulations, our algorithm starts with an equally high accuracy rate of 91.2583% for single dark instances. Its accuracy does dwindle more as we require longer chains, but still reaches an accuracy of 66.9014% for long dark sequences (chain length 15).

Overall, we observe that our method detects chains of the same type of influence with very high accuracy, implying it is ready to be deployed in real-world applications (see following section) where the requirements for consecutive predictions may not be as demanding as we have tested for here.

Table 4.1: Results on synthetic interdependencies data

Col. ID	**A**	**B**	**C**	**D**	**E**	**F**
Chain	Synthetic Positive Influence % (over all simulations), κ="positive", chance: 90%	Positive Influence % detected (over Col. A)	Synthetic Negative Influence % (over all simulations), κ="negative", chance: 90%	Negative Influence % detected (over Col. C)	Synthetic Dark Influence % (over all simulations), κ="dark", chance: 90%	Dark Influence % detected (over Col. E)
1	0.9031	0.9356	0.9032	0.9344	0.9037	0.9125
2	0.8574	0.9193	0.8574	0.9179	0.8569	0.8701
3	0.8136	0.9030	0.8139	0.9063	0.8144	0.8276
4	0.7725	0.8956	0.7735	0.8966	0.7735	0.8010
5	0.7339	0.9003	0.7342	0.8978	0.7338	0.7679
6	0.6973	0.8875	0.6968	0.8927	0.6962	0.7239
7	0.6619	0.8907	0.6612	0.8818	0.6603	0.7232
8	0.6291	0.8777	0.6286	0.8743	0.6273	0.6414
9	0.5974	0.8872	0.5962	0.8802	0.5962	0.6640
10	0.5669	0.8739	0.5667	0.8731	0.5659	0.6192
11	0.5387	0.8640	0.5386	0.8539	0.5377	0.6690
12	0.5127	0.8455	0.5123	0.8624	0.5109	0.6419
13	0.4865	0.8461	0.4857	0.8503	0.4854	0.6250
14	0.4615	0.8426	0.4612	0.8471	0.4613	0.5312
15	0.4387	0.8294	0.4379	0.8221	0.4391	0.5294

4.8 Supplementary Information for the Applications

4.8.1 Details of the desert ecosystem

In order to reconstruct the ecosystem's causal network, we took the common time span for which data were available in all four categories, namely, from 1993 to 2009. We implemented our method on time series constructed from 32 data inputs (Table S2), with $E = 2, \tau = 1$ and $h = 0$. We used $h = 0$ because the quantities were averaged to a yearly basis from varying time frequencies. Thus, for this application, we wanted to examine the concurrent interactions in the system from year to year and not the predictions. To keep the strongest interactions, we filtered the network links with the "Maximum Spanning Tree" algorithm (Hu 1961) using the weights retrieved from Eq. (4.11) of our method. As a result, we plotted six periods of interest (Figure 4.1) in order to depict the ecosystem's synergies.

Table 4.2: Variables and species in the Chihuahuan Desert ecosystem in Arizona

Weather	**Rodents**	**Plants**	**Ants**
Temperature	*Ammospermophilus harrisi*	*Astragalus nuttallianus*	*Conomyrma bicolor*
Precipitation	*Dipodomys merriami*	*Cassia bauhinoides*	*Conomyrma insane*
	Dipodomys ordii	*Ephedra trifurca*	*Iridomyrma pruinosum*
	Neotoma albigula	*Eriastrum diffusum*	*Novomessor cockerelli*
	Onychomys torridus	*Erodium cicutarium*	*Pheidole sitarches*
	Peromyscus eremicus	*Gutierrezia sarothrae*	*Pheidole xerophila*
	Chaetodipus penicillatus	*Lesquerella gordoni*	*Pogonomyrmex desertorum*
	Reithrodontomys megalotis	*Perezia nana*	*Solenopsis sp.*
	Spermophilus spilosoma	*Plantago patagonica*	*Solenopsis xyloni*
		Solanum eleagnifolium	
		Phemeranthus angustissimum	
		Phemeranthus aurantiacus	

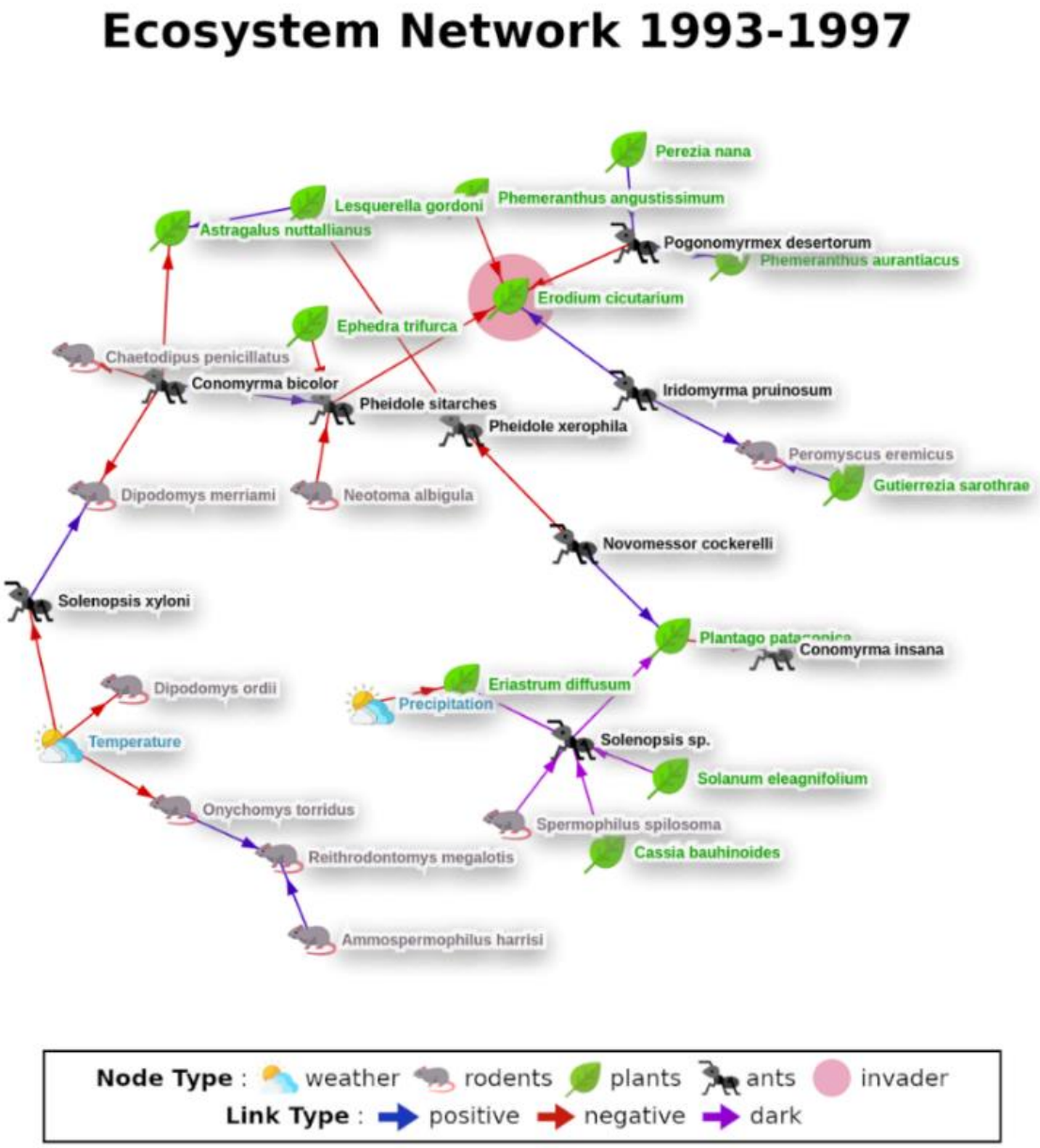


Figure 4.4. Aggregate causal network before the "aggression" of Erodium cicutarium. The node icon is representative of the node's type (ants, plants, rodents, weather). The link color denotes type of causality (blue for positive, red for negative and purple for dark).

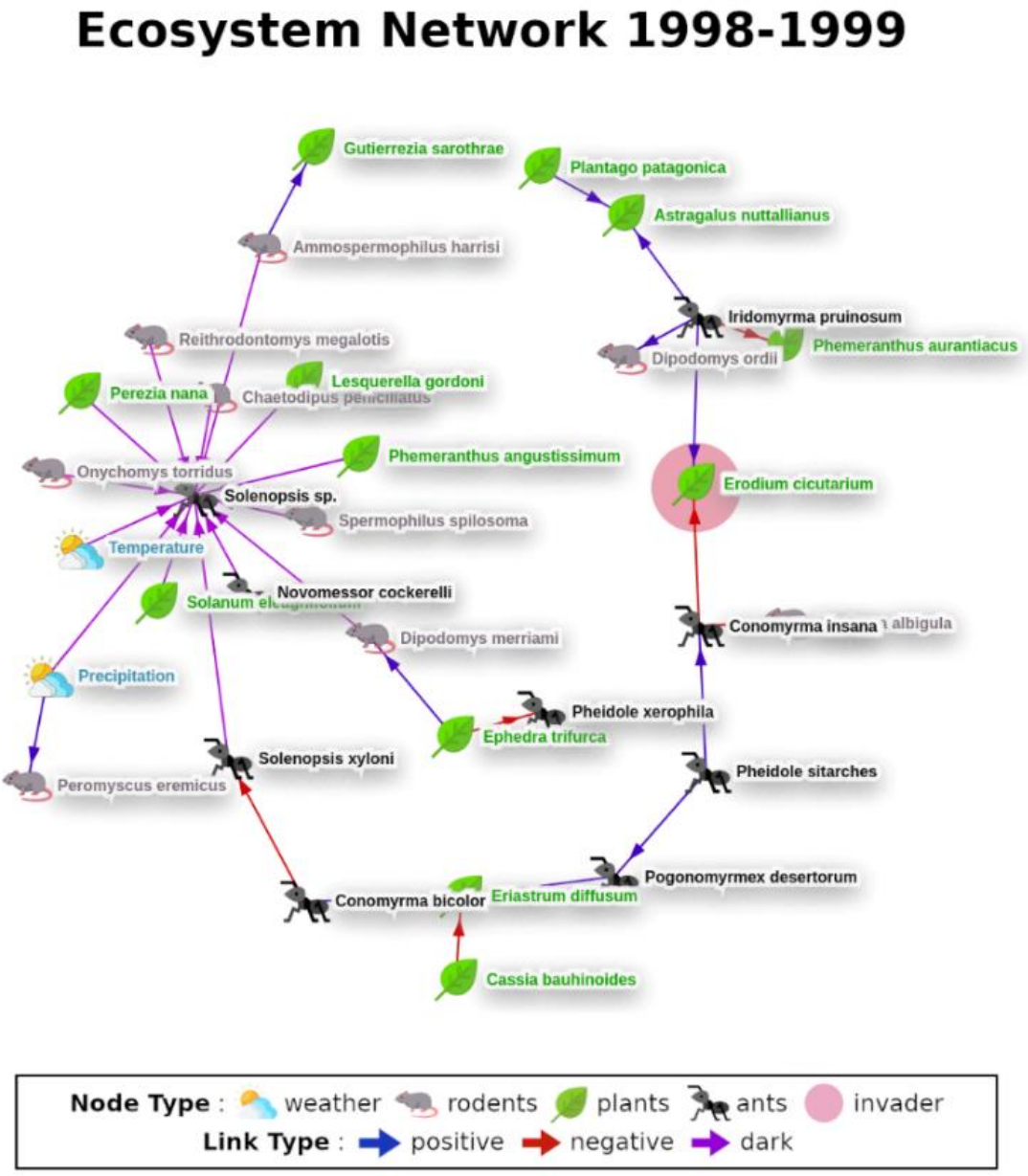


Figure 4.5. Aggregate causal network during the invasion period of Erodium cicutarium. The node icon is representative of the node's type (ants, plants, rodents, weather). The link color denotes type of causality (blue for positive, red for negative and purple for dark).

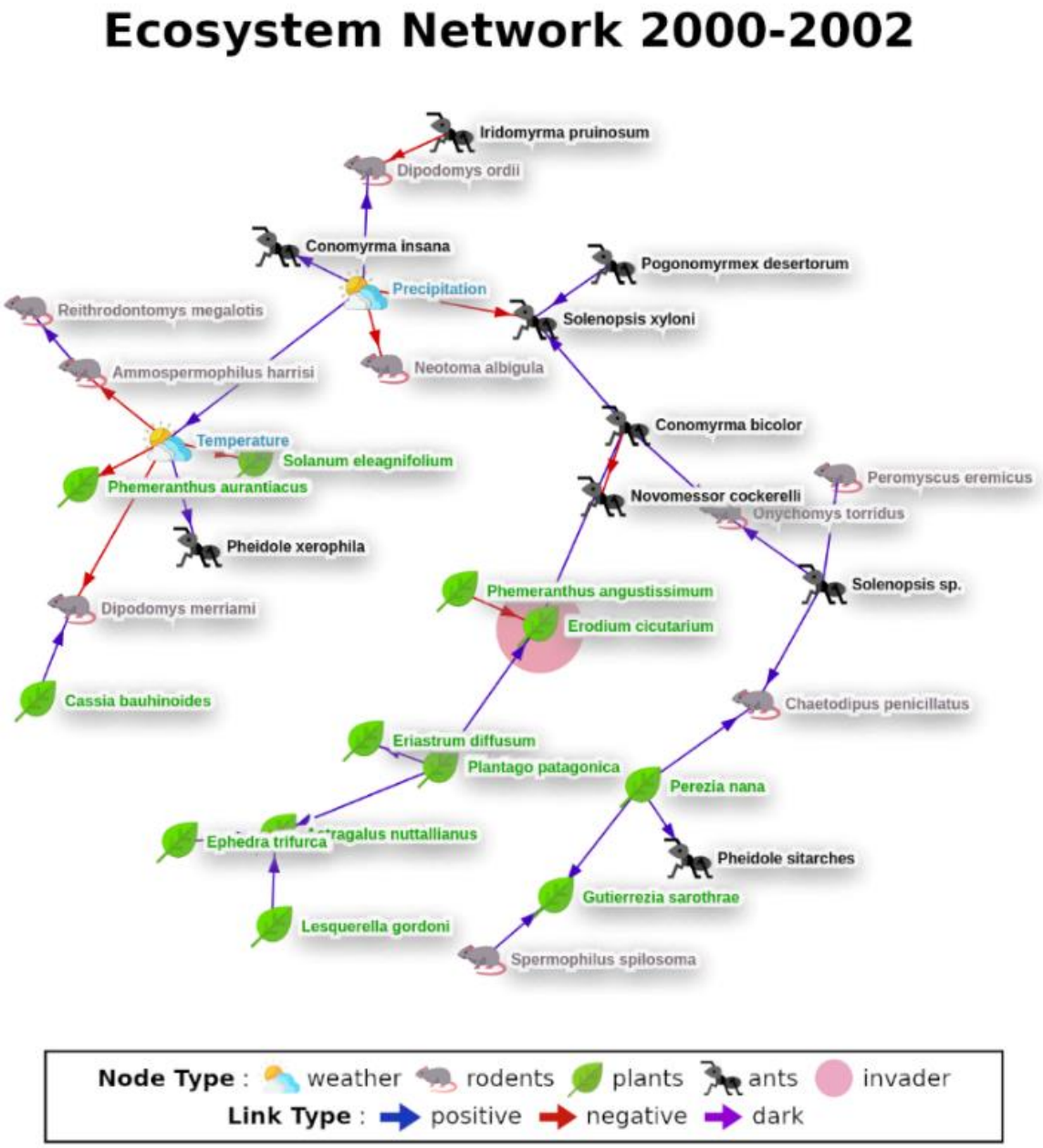


Figure 4.6. Aggregate causal network during the invasion period of Erodium cicutarium. The node icon is representative of the node's type (ants, plants, rodents, weather). The link color denotes type of causality (blue for positive, red for negative and purple for dark).

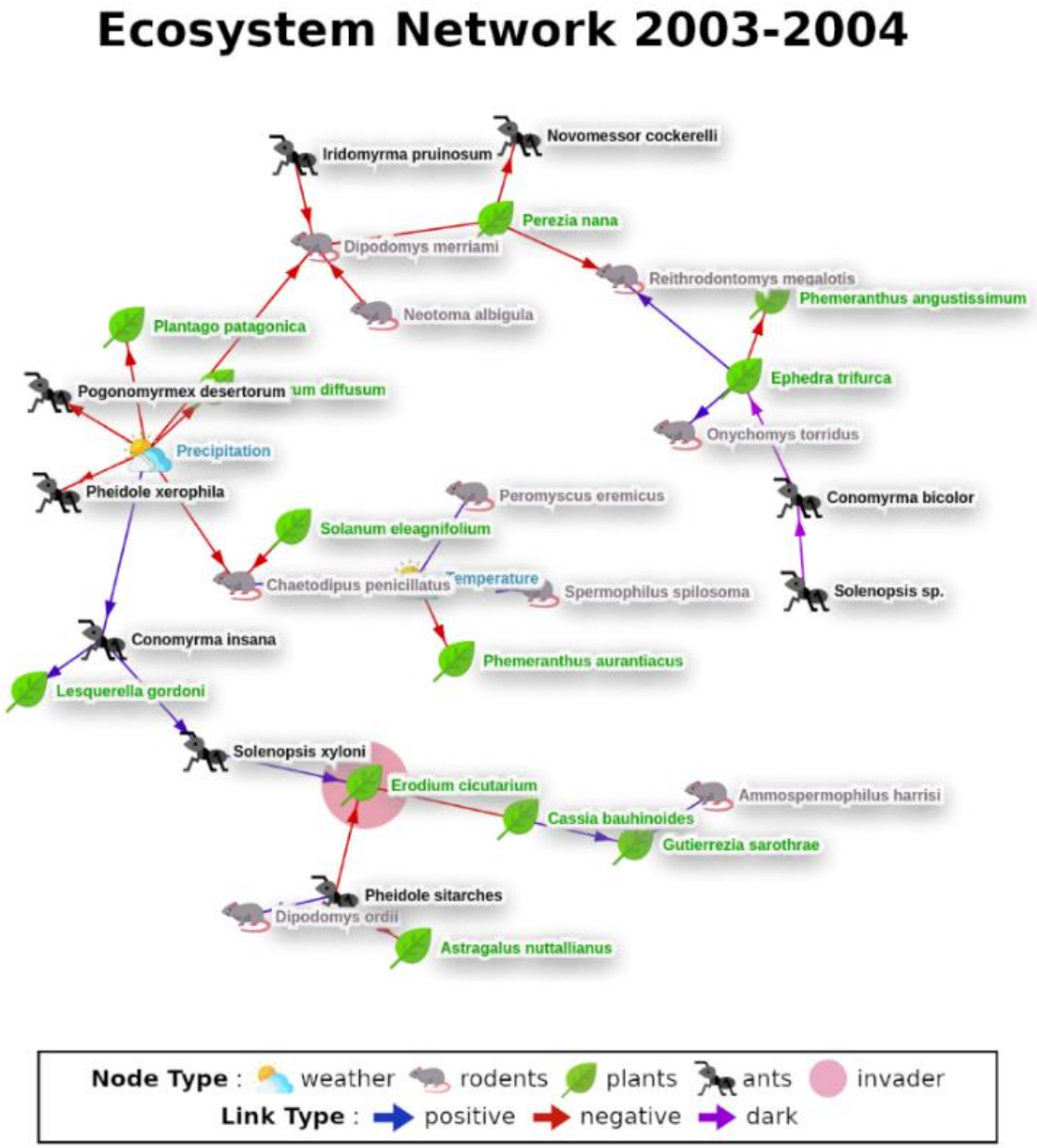


Figure 4.7. Aggregate causal network during the invasion period of Erodium cicutarium. The node icon is representative of the node's type (ants, plants, rodents, weather). The link color denotes type of causality (blue for positive, red for negative and purple for dark).

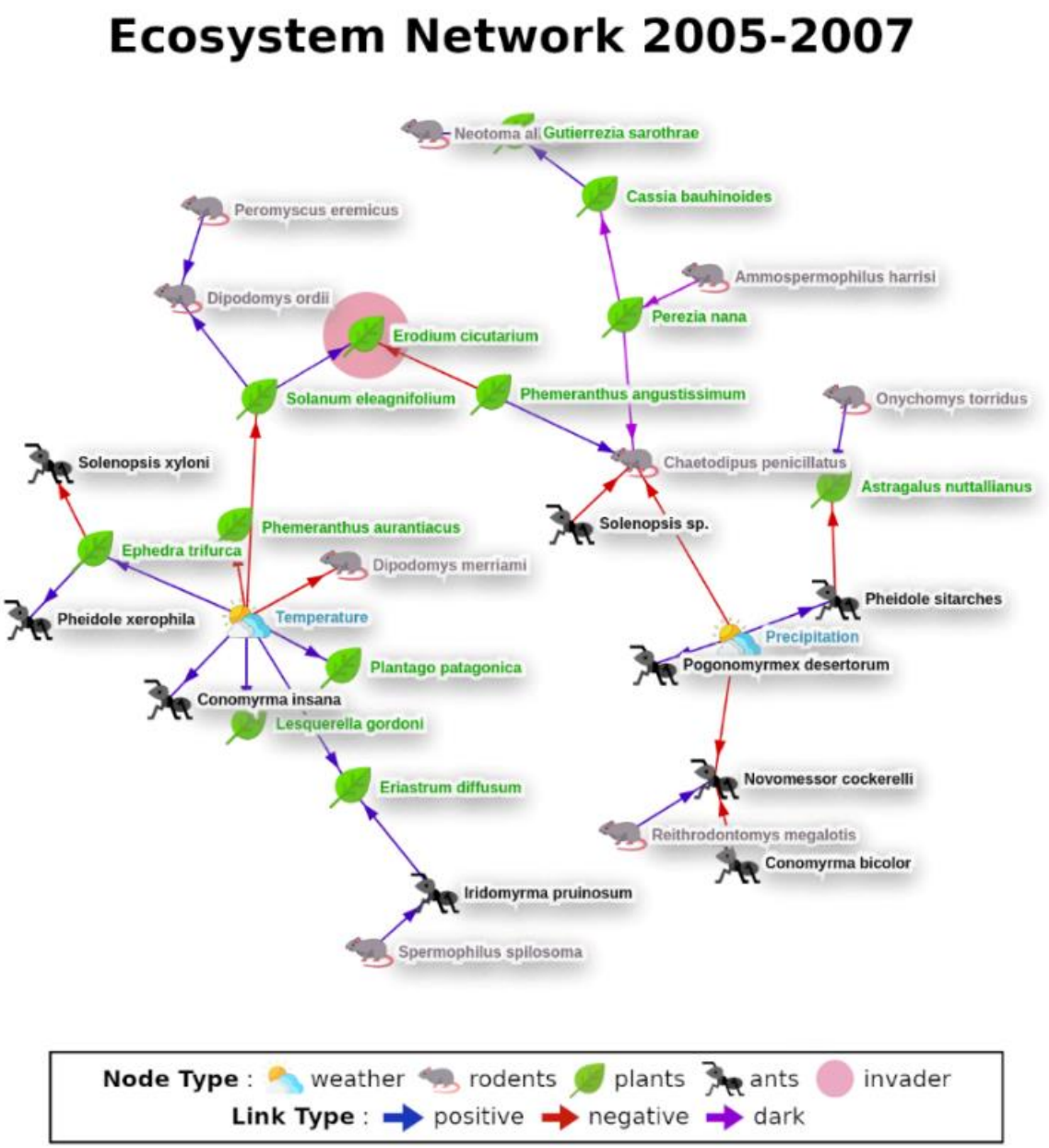


Figure 4.8. Aggregate causal network during the invasion period of Erodium cicutarium. The node icon is representative of the node's type (ants, plants, rodents, weather). The link color denotes type of causality (blue for positive, red for negative and purple for dark).

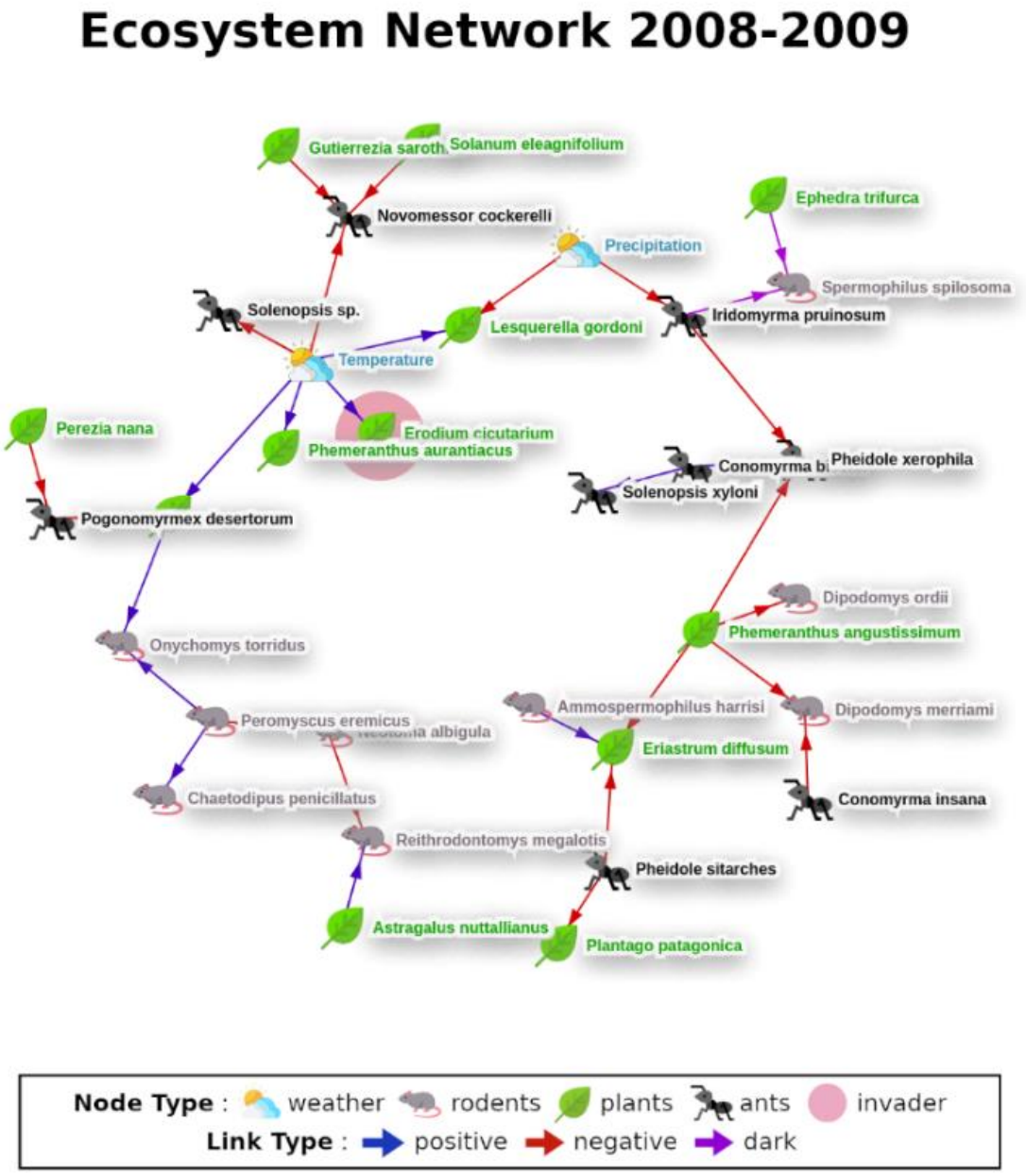


Figure 4.9. Aggregate causal network during the post-invasion period of Erodium cicutarium. The node icon is representative of the node's type (ants, plants, rodents, weather). The link color denotes type of causality (blue for positive, red for negative and purple for dark).

4.8.2 Details of brain system

For each of the 20 subjects, we had five trials and for each trial we had time series (L=256) from the 64 electrodes' voltage recordings. In order to have a panoramic comparison of EEG activity between the alcoholic and control subjects we adhered to a meticulous procedure. First, for each subject and for each trial, we calculated the underlying EEG network using the 64 electrodes as nodes and $E = 3, \tau = 1$ and $h = 1$ as parameters. At this point, we had 20×5×3=300 networks (three corresponding to the positive, negative and dark aspects). Then, for each subject, we created three resultant dynamic networks, averaging across the five trials. At this point, we had 20×3=60 averaged dynamic networks accounting for all 20 subjects. Finally, in order to compare the alcoholic versus control network structures, we averaged out the 10×3 networks for each type of subject (alcoholic and control). Thus, for our analysis, we kept three networks (positive, negative and dark) for the "average alcoholic" and three for the "average control" brain. When we refer to three networks, in fact, we are viewing the same network from three different aspects (positive, negative and dark), for the alcoholic and control subjects separately.

Table 4.3: List of all electrodes according to brain region

Frontal region (1)	**Central region (2)**	**Parietal region (3)**	**Occipital region (4)**	**Temporal region (5)**	**Auxiliary electrodes (*)**
FP1	*FC1*	*CP1*	*PO1*	*T7*	*X*
FP2	*FC2*	*CP2*	*PO2*	*T8*	*Y*
FPZ	*FC3*	*CPZ*	*POZ*	*TP7*	*nd*
AF7	*FC4*	*CP3*	*PO7*	*TP8*	
AF8	*FC5*	*CP4*	*PO8*	*CP5*	
AF1	*FC6*	*PZ*	*O1*	*CP6*	
AF2	*FCZ*	*P1*	*O2*	*P7*	
AFZ	*CZ*	*P2*	*OZ*	*P8*	
F7	*C1*	*P3*		*P5*	
F8	*C2*	*P4*		*P6*	
F6	*C3*			*FT7*	
F3	*C4*			*FT8*	
F4	*C5*				
F1	*C6*				
F2					
FZ					
F5					

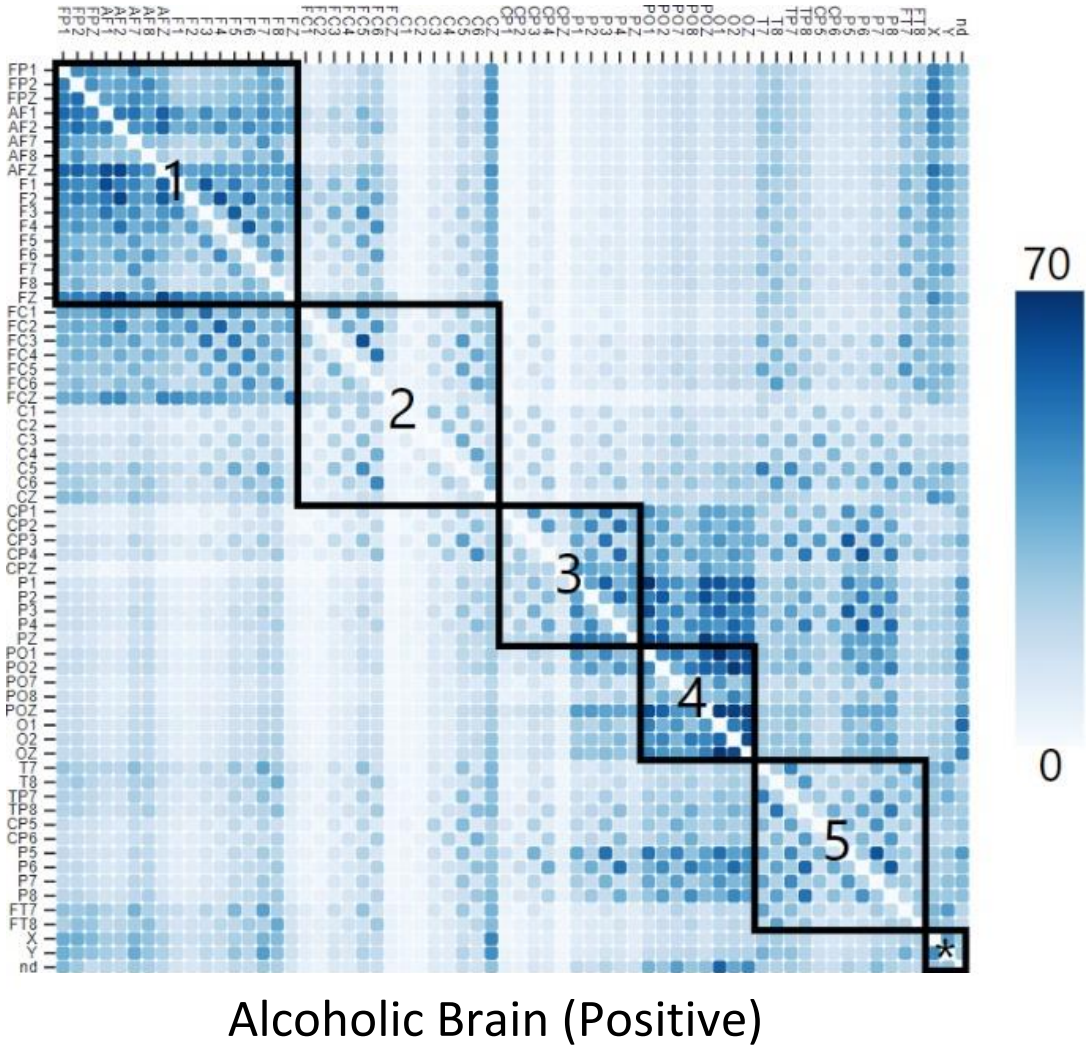


Figure 4.10. Cumulative adjacency matrix for the average positive network structure of alcoholic brain. Box 1 corresponds to frontal region, box 2 corresponds to central region, box 3 corresponds to parietal region, box 4 corresponds to occipital region, box 5 corresponds to temporal region, and box * concerns auxiliary electrodes.

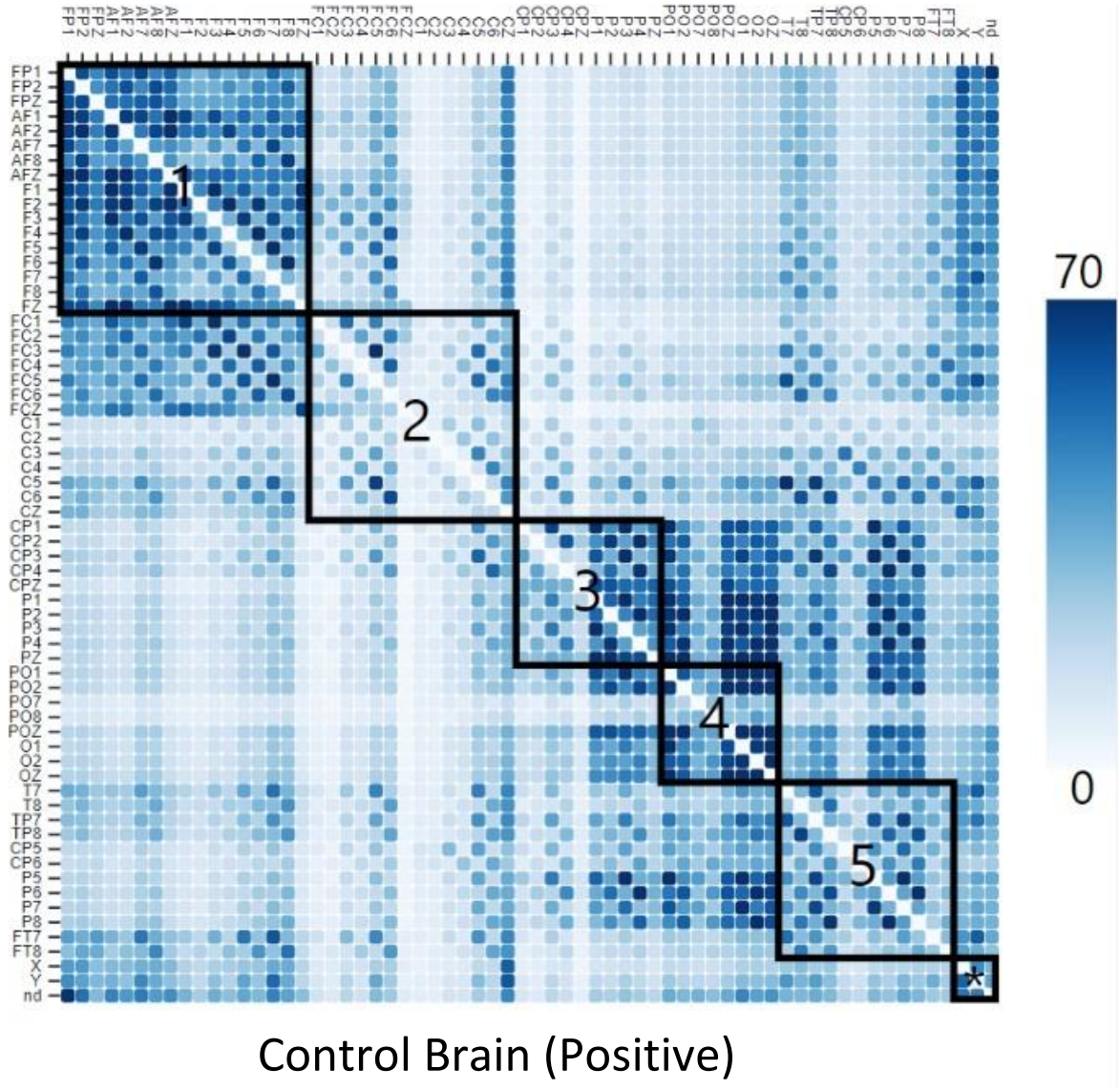
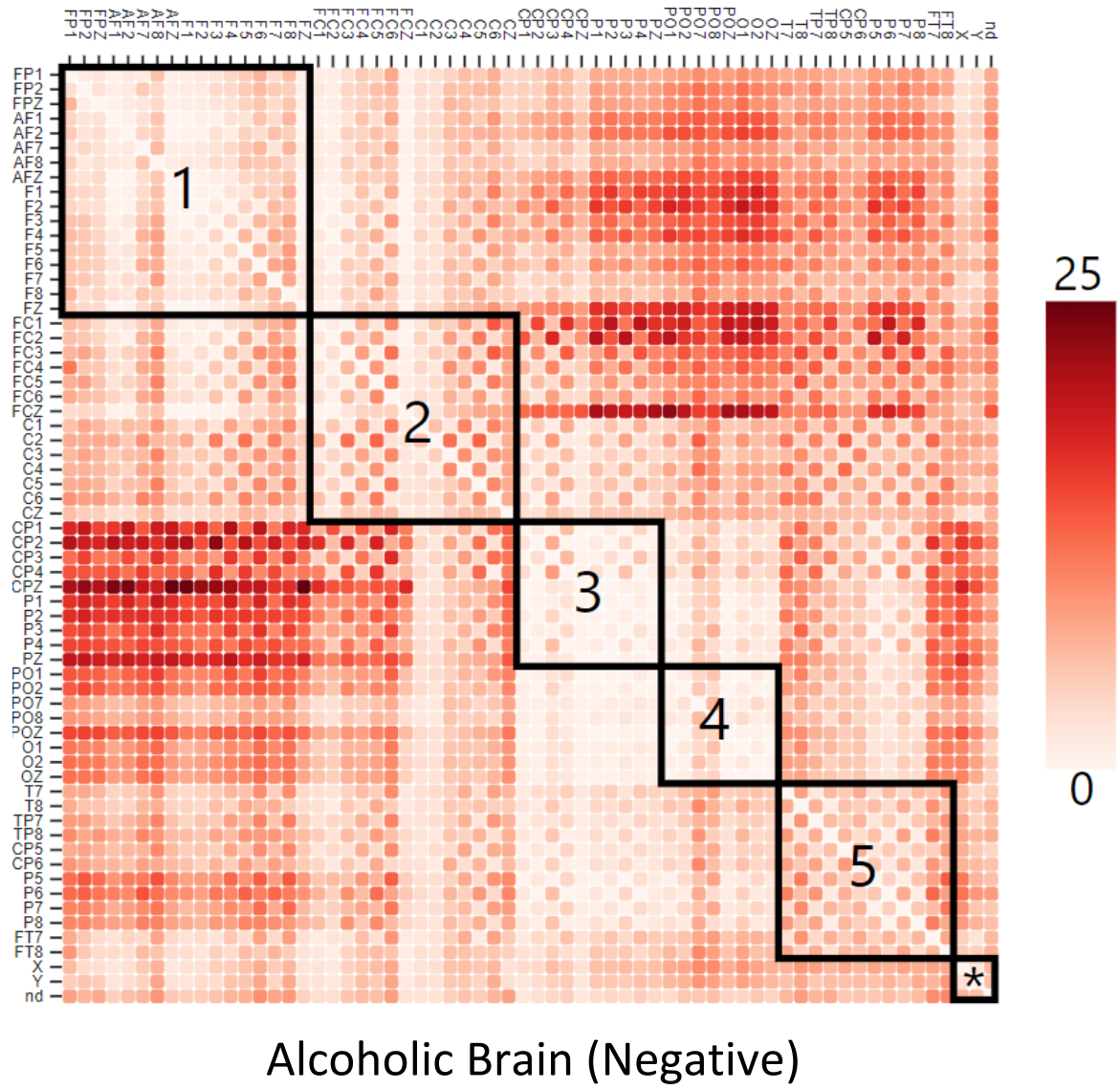


Figure 4.11. Cumulative adjacency matrix for the average positive network structure of control brain. Box 1 corresponds to frontal region, box 2 corresponds to central region, box 3 corresponds to parietal region, box 4 corresponds to occipital region, box 5 corresponds to temporal region, and box * concerns auxiliary electrodes.



Alcoholic Brain (Negative)

Figure 4.12. Cumulative adjacency matrix for the average negative network structure of alcoholic brain. Box 1 corresponds to frontal region, box 2 corresponds to central region, box 3 corresponds to parietal region, box 4 corresponds to occipital region, box 5 corresponds to temporal region, and box * concerns auxiliary electrodes.

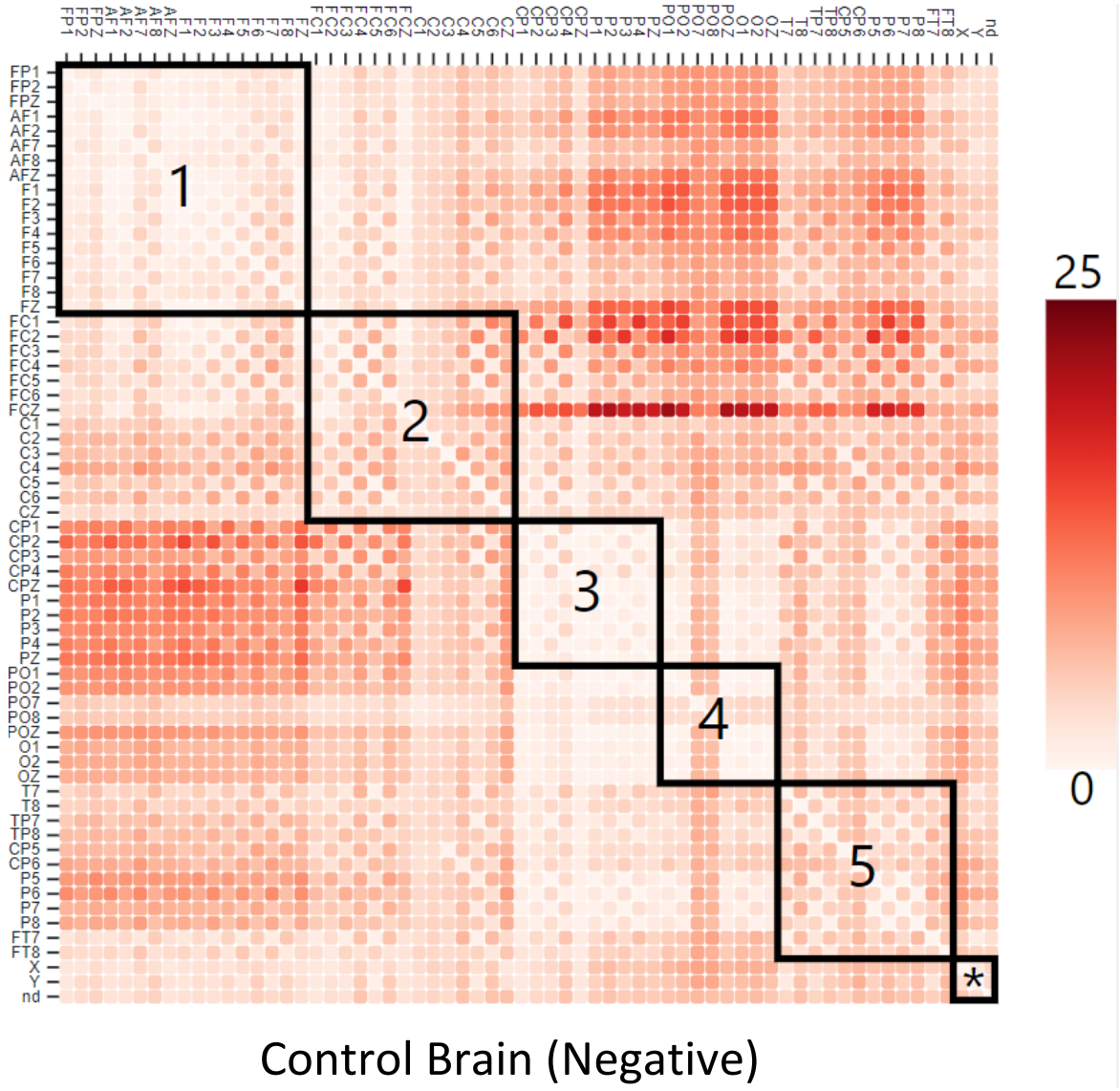
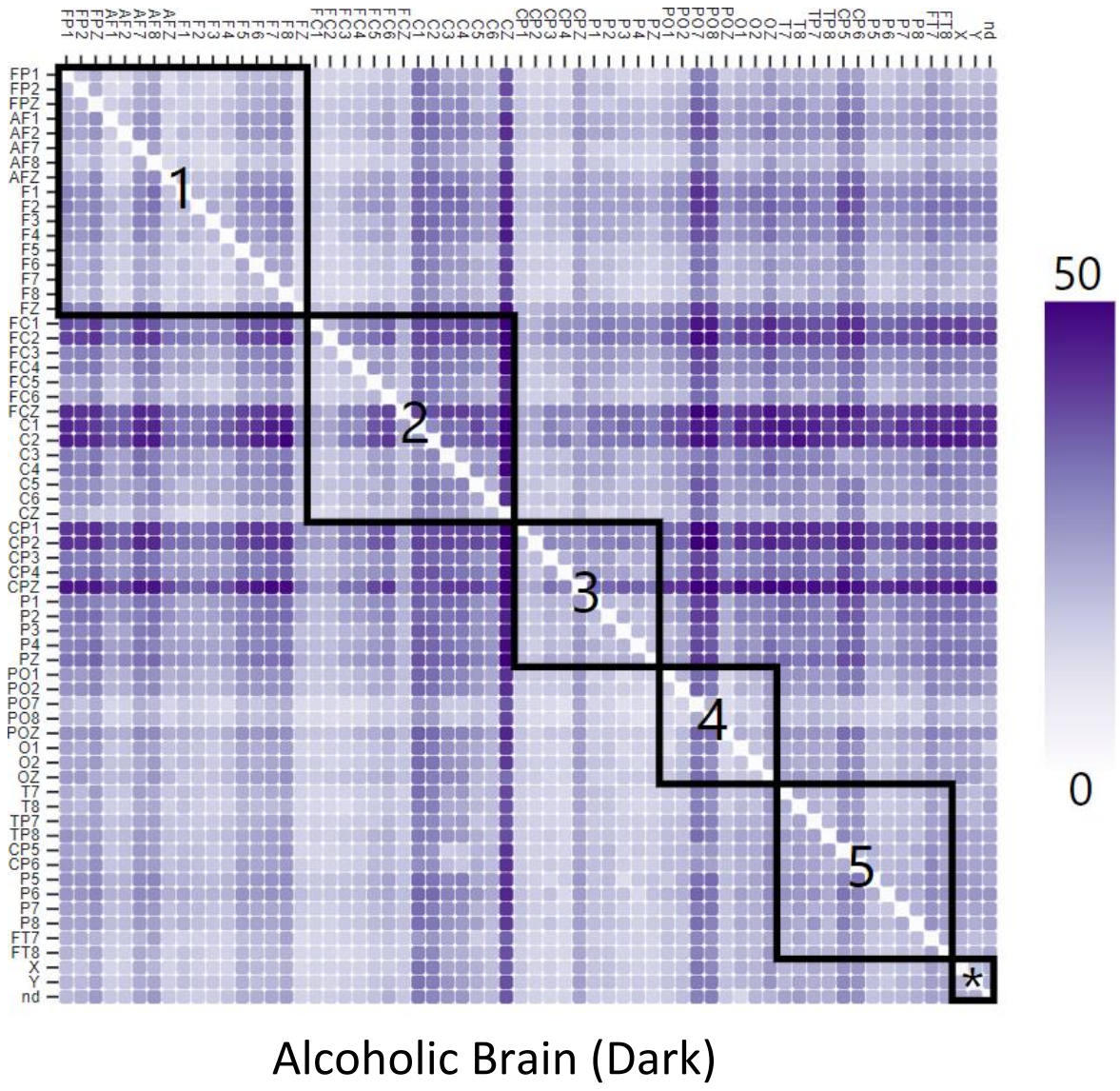


Figure 4.13. Cumulative adjacency matrix for the average negative network structure of control brain. Box 1 corresponds to frontal region, box 2 corresponds to central region, box 3 corresponds to parietal region, box 4 corresponds to occipital region, box 5 corresponds to temporal region, and box * concerns auxiliary electrodes.



Alcoholic Brain (Dark)

Figure 4.14. Cumulative adjacency matrix for the average dark network structure of alcoholic brain. Box 1 corresponds to frontal region, box 2 corresponds to central region, box 3 corresponds to parietal region, box 4 corresponds to occipital region, box 5 corresponds to temporal region, and box * concerns auxiliary electrodes.

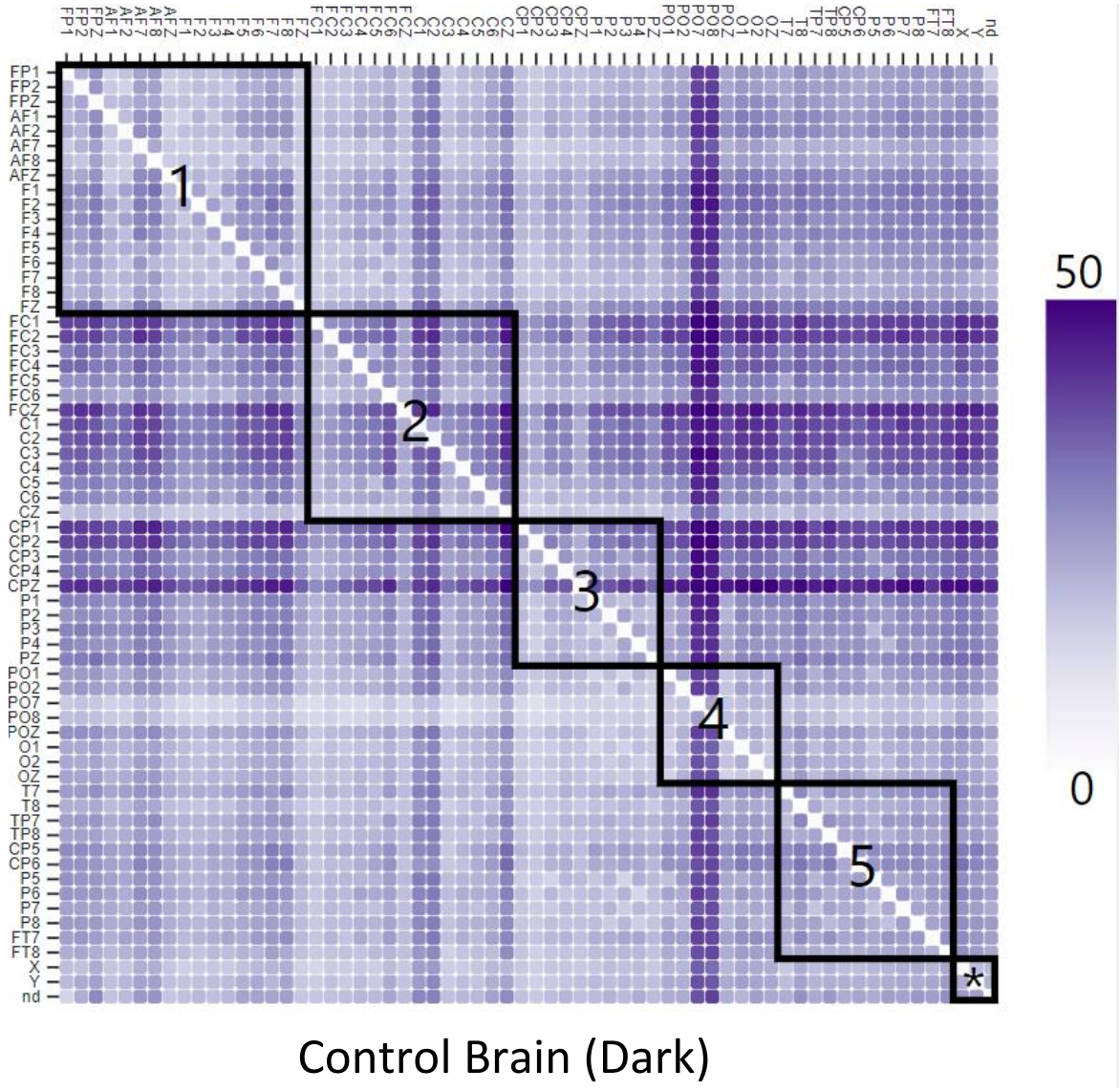


Figure 4.15. Cumulative adjacency matrix for the average dark network structure of control brain. Box 1 corresponds to frontal region, box 2 corresponds to central region, box 3 corresponds to parietal region, box 4 corresponds to occipital region, box 5 corresponds to temporal region, and box * concerns auxiliary electrodes.

4.8.3 Details of CDS network

Initially, we built a next-day-prediction network (day by day) from 2007 December 14 to 2019 May 13. To do this, we extracted the underlying interdependencies, with $E = 3, \tau = 1$ and $h = 1$. Thus, we could view the same network from three different spectra (positive, negative and dark). In Figures 4.16-4.18, we present a cumulative adjacency matrix for each spectrum, where the color is associated with cumulative intensity. Darker colors denote stronger links overall.

Table 4.4: Senior prime banking Credit Default Swaps (CDS) of five-year maturity

1	Alliance & Leicester International Limited	35	Deutsche Bank AG
2	Allied Irish Banks	36	DNB ASA
3	Alpha Bank SA	37	Erste Group Bank AG
4	American Express Co	38	Fun Banc Caixa Dest
5	Australia and New Zealand Banking Group	39	Goldman Sachs Group
6	Banca Italease SPA	40	HSBC Bank PLC
7	Banca Monte Paschi	41	HSBC Holdings PLC
8	Banco Com Portugues	42	IKB Deutsche Industriebank AG
9	Banco De Sabadell	43	ING Bank N.V.
10	Banco Pop Espanol	44	Intesa Sanpaolo SPA
11	Banco Santander SA	45	Irish Bank Resolution Corp.
12	Bank of America	46	KBC Bank
13	Bank of Ireland	47	Landesbank Baden-Württemberg
14	Bank of Scotland	48	Macquarie Bank Ltd
15	Barclays Bank PLC	49	Mediobanca SPA
16	Barclays SLCSM Limited	50	National Bank of Greece SA
17	Bawag PSK	51	National Australian Bank
18	Bayerische Landesbank	52	Natixis
19	Banco Bilbao Vizcaya Argentaria SA	53	Natwest Markets PLC
20	Banca Nazionale Del Lavoro	54	Norddeutsche Landesbank
21	Bank of New York Mellon Corp.	55	Nordea Bank AB
22	Bank of America NA	56	Portigon AG
23	Banco Espírito Santo	57	Raif Zentralbank
24	BNP Paribas SA	58	Santander UK PLC
25	Capital One Bank USA NA	59	Skandinaviska Enskilda Banken
26	Capital One Financial	60	Societe Generale
27	CDA De Valencia Castellan	61	Standard Chartered Bank
28	CDA Y MP De Madrid	62	Skandinaviska Enskilda Banken
29	Citigroup Inc.	63	The PNC Financial Services Group
30	CMWL Bank Of Australia	64	The Royal Bank of Scotland Group plc
31	Commerzbank AG	65	UBS AG
32	Coop Rabobank UA	66	Unione Di Banche
33	Danske Bank A/S	67	Van Lanschot N.V.
34	De Volksbank NV	68	Westpac Banking Corp

Table 4.5: Top 10 most influential CDS of all time, ranked by subtracting cumulative in-strength centrality from out-strength centrality

Positive (Top 10 CDS)	**Negative (Top 10 CDS)**	**Dark (Top 10 CDS)**
Skandinaviska Enskilda Banken	Landesbank Baden-Württemberg	Santander UK PLC
Nordea Bank AB	Bawag PSK	IKB Dt Indstrbk AG
Skandinaviska Enskilda Banken	IKB Dt Indstrbk AG	Capital One Financial
Coop Rabobank UA	Norddeutsche Landesbank	Skandinaviska Enskilda Banken
Landesbank Baden-Württemberg	Alpha Bank SA	American Express Co
Danske Bank A/S	The PNC Financial Services Group	Capital One Bank USA NA
Bayerische Landesbk	National Bank Of Greece SA	Bawag PSK
HSBC Holdings PLC	Barclays SLCSM Ltd	Bayerische Landesbk
HSBC Bank PLC	Portigon AG	Cmwl Bk of Australia
Capital One Bank USA NA	Van Lanschot N.V.	Australia and New Zealand Banking Group

Figure 4.4: Cumulative Adjacency Matrix for the positive type of interdependencies on the last day of the time series data (2019 May 13). Darker color denotes higher accumulated link strength (forecasting).

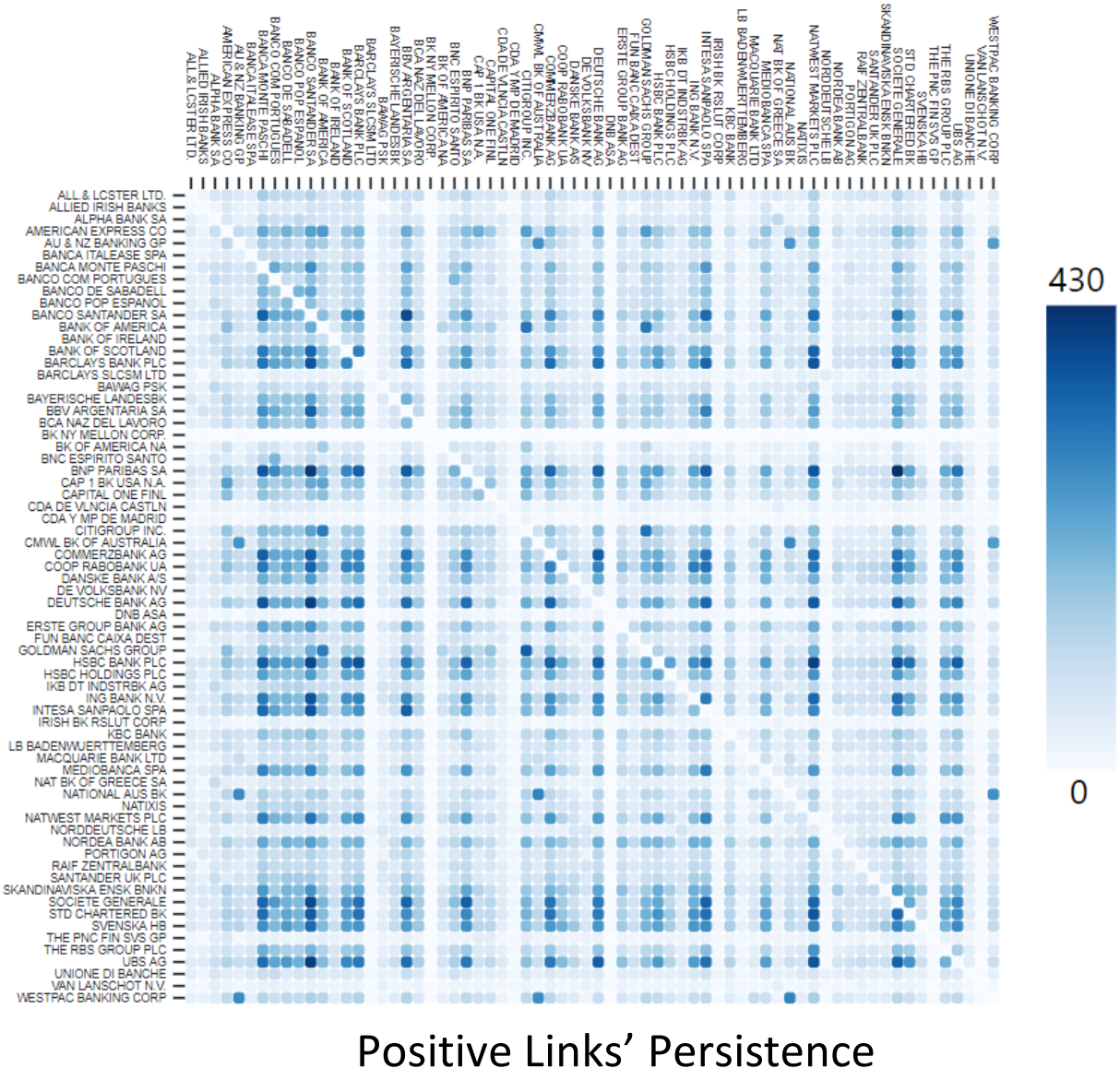
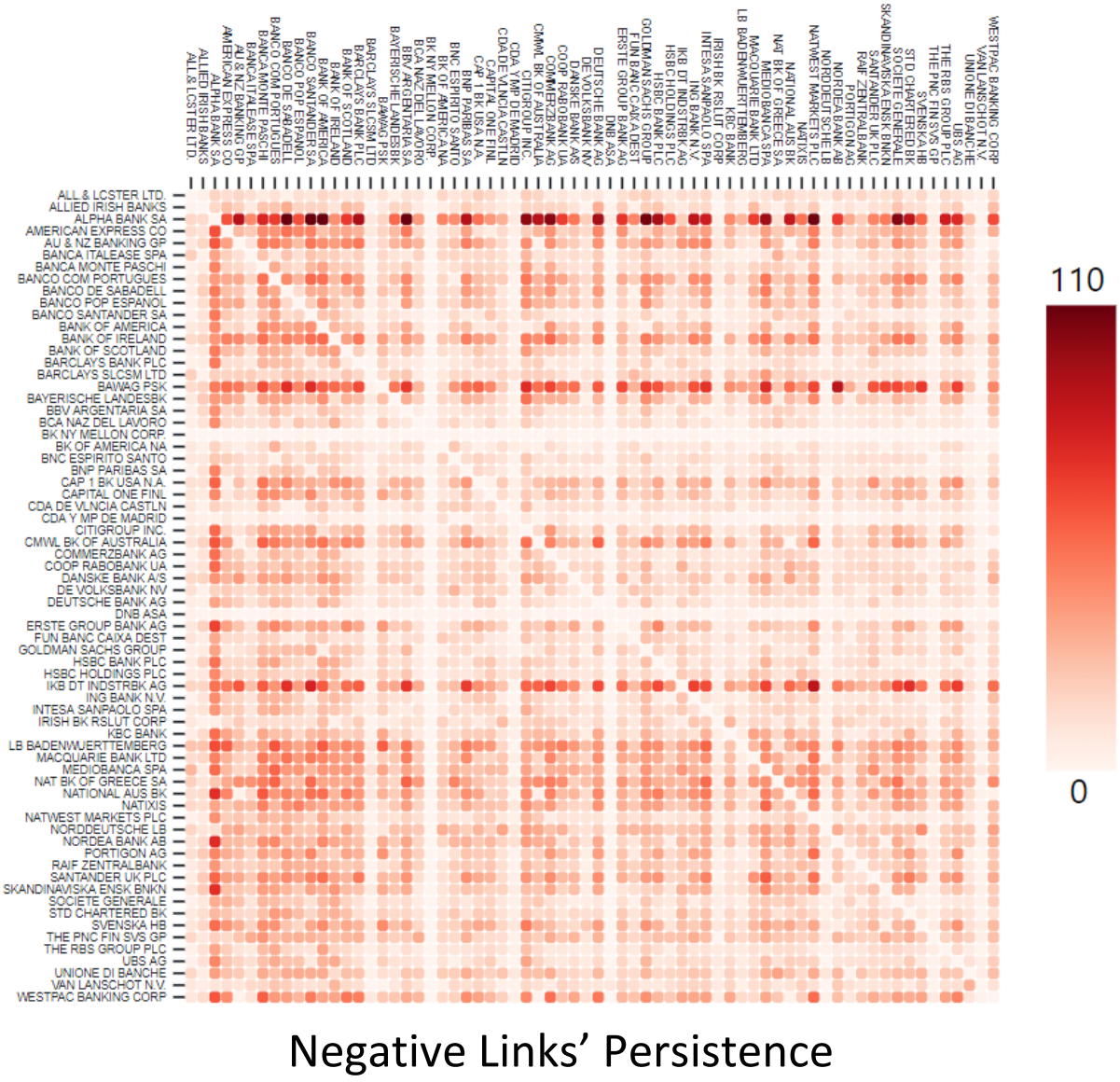
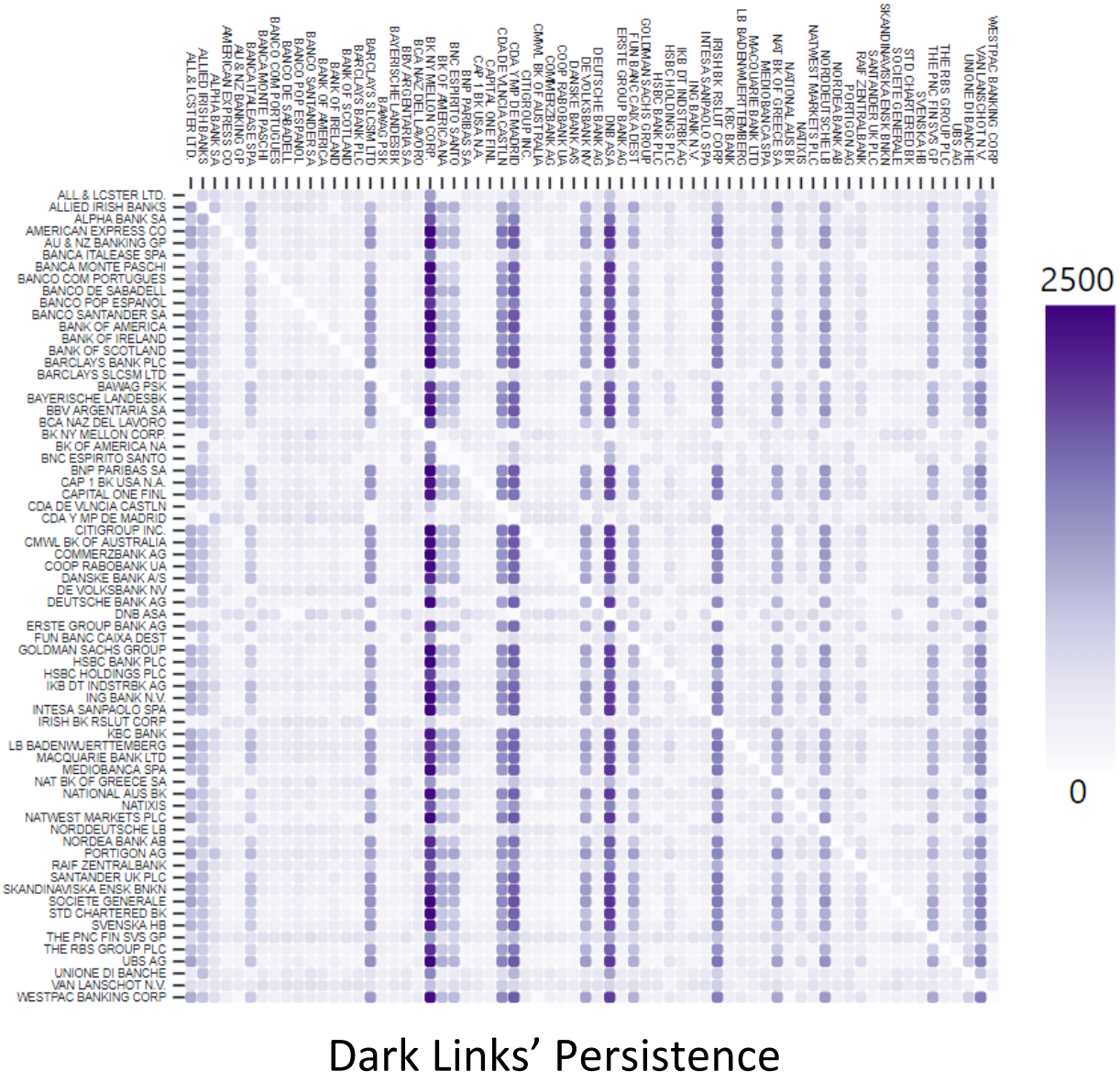


Figure 4.5: Cumulative Adjacency Matrix for the negative type of interdependencies on the last day of the time series data (2019 May 13). Darker color denotes higher accumulated link strength (forecasting).



Negative Links' Persistence

Figure 4.6: Cumulative Adjacency Matrix for the dark type of interdependencies on the last day of the time series data (2019 May 13). Darker color denotes higher accumulated link strength (forecasting).



4.9 From Time Series to Reconstructed Attractors

Dynamical systems theory states that the temporal evolution of a system is defined in some state space (or phase space, for continuous systems). The perception of the "state of a system" is powerful, even for nondeterministic systems. For example, stochastic Markov processes can be expressed through a set of states, along with a set of transition probabilities which define the random transition rules of the system. Specifying a point ω in this space $\mathbb{R}^m$ specifies the state of the system and vice versa. Thus, in this paper, in order to tap into the dynamics of the system, we employ the dynamics of the points in the corresponding state space, through an m-dimensional mapping. A sequence of points $\omega(t)$ is called a "trajectory" of the dynamical system, and $\omega(0)$ denotes the respective "initial conditions". For many dynamical systems, the trajectory will, after some time, be attracted to some subset of the state space. This set is invariant under the dynamical evolution, is called the "attractor" of the system, and can be studied as a manifold (Kantz & Schreiber 2004).

At their conception, time-series analysis and manifolds were two remotely distant areas of research, the first being on the very practical edge of statistics, and the second, antithetically, in the sphere of pure mathematics. A first promising result with the potential to bridge the two seemingly incompatible disciplines was Whitney's embedding theorem (1936), which suggests that a generic map from an *m*-dimensional manifold *M* to $\mathbb{R}^{2m+1}$ is an embedding, i.e. the image of *M* is completely unfolded in the higher-dimensional space. Notably, no two points in *M* map to the same point in $\mathbb{R}^{2m+1}$ (injective property). As $2m+1$ independent measurements (a time series) of a system can be considered a map from the set of states of *M* to $\mathbb{R}^{2m+1}$, Whitney's theorem suggests that each state can be identified uniquely by a vector of $2m+1$ time-series measurements, thereby reconstructing the state space. However, practically speaking, most scientists end up with some time series without any awareness of the overall state space and its dimension m. Moreover, even if they did know the value of m, according to Whitney, they would need $2m+1$ distinct time series to be able to reconstruct the original state space. Due to the apparent impracticality for experimental settings, Whitney's theorem, despite being monumental for differential topology, did not ignite a connection between time series and manifolds.

Half a century later, Takens' (1981) embedding theorem, along with a first practical study (Packard et al., 1980), bridged the gap, and a burgeoning literature of time-series methods capitalized on the newly unlocked insights from differentiable manifolds. The contribution of

Takens' theorem was to show that state space reconstruction could be achieved with just a single time series. Takens proved that, instead of $2m + 1$ distinct time series, the time-delayed versions $[X(t), X(t-\tau), X(t-2\tau), \dots, X(t-2m\tau)]$ of one time series X would suffice to embed the m-dimensional manifold. Moreover, Takens showed that reconstruction is viable even in a space with dimension $E \leq 2m + 1$. This theorem liberated the attractor reconstruction task from the need to find $2m + 1$ distinct time series and enabled every single time series to be analyzed in its time-delayed form, provided that m could somehow be approached. A decade later, Sauer et al. (1991) were able to generalize both Whitney's and Takens' theorems to fractal dimensions as well.

Therefore, practically speaking, in order to study the hidden dynamics of a time series, one needs to embed it in a dimension E using a lag τ. In order to retrieve the optimal combination of E and τ, there are various methods in the literature (Fraser & Swinney 1986; Farmer & Sidorowich 1987; Casdagli 1989; Liebert & Schuster 1989; Kennel et al. 1992; Cellucci et al. 2003; Krakovska et al. 2015). Nevertheless, with today's computational power, optimal pairs of E and τ can be found by trial and error, comparing the results in terms of forecasting accuracy for a whole range of reasonable embedding parameters (E, τ).

4.10 Theorems and Proofs

Let X, Y be any two variables evolving through time which belong to a common dynamical system. Their trajectories $X(t)$ and $Y(t)$ are attracted over time to the m-dimensional attractor manifold M of the system. Let the time lags of $X(t)$ and $Y(t)$ be embedded into a state space with dimension E and, using delay τ, let them create E-dimensional trajectories of vectors $\underline{x}(t) =< X(t), X(t-\tau), \dots, X(t-(E-1)\tau) >$ and $\underline{y}(t) =< Y(t), Y(t-\tau), \dots, Y(t-(E-1)\tau) >$ respectively. The sets of all points $\underline{x}(t)$ and $\underline{y}(t)$ are known as reconstructed attractor manifolds or shadow attractors and are notated as M_X and M_Y respectively. Since causality (statistical influence) is the quantifiable influence exerted by one variable X on another variable Y, unlike correlation which distinguishes between positive and negative relations in time-series analysis, the other measures of statistical influence tend to ignore the nature of interactions (for more discussion, see Stavroglou et al. 2019). Therefore, as a next step, an important definition is provided, which enables identification of the nature of influence.

Definition 4.1.

(1) If X causes same-direction changes to Y, then we say that X has a positive influence on Y.

(2) If X causes opposite-direction changes to Y, then we say that X has a negative influence on Y.

(3) If X causes changes to Y which are of neither the same nor the opposite direction, then we say that X has a "dark" influence on Y.

The following results provide the mathematical setting needed for deducing influence in a dynamical systems framework. In what follows, by smooth functions, we mean at least $\mathbb{C}^2$ (the derivatives of up to order 2 exist and are continuous).

Lemma 4.1. *Let M be an m-dimensional compact manifold and $X: M \to \mathbb{R}$, $Y: M \to \mathbb{R}$ be smooth observation functions. Let $\varphi: M \rightarrowtail M$ be a smooth diffeomorphism. If there exists $\psi: M_X \rightarrowtail M_Y$ such that ψ is bijective, then M_X causes M_Y.*

Proof of Lemma 4.1: Since X, Y are smooth functions and φ is a smooth diffeomorphism, according to Takens' theorem it is a generic property that the maps $\Phi_{(\varphi,X)}(\underline{\omega}): M \to \mathbb{R}^{2m+1}$ and $\Phi_{(\varphi,Y)}(\underline{\omega}): M \to \mathbb{R}^{2m+1}$, defined as

$$\Phi_{(\varphi,X)}(\underline{\omega}) = \langle X(\underline{\omega}), X(\varphi(\underline{\omega})), \dots, X(\varphi^{2m}(\underline{\omega})) \rangle$$

$$\Phi_{(\varphi,Y)}(\underline{\omega}) = \langle Y(\underline{\omega}), Y(\varphi(\underline{\omega})), \dots, Y(\varphi^{2m}(\underline{\omega})) \rangle$$

are embeddings (reconstructions) of the original manifold M and $\underline{\omega} \in M$.

$$\text{Specifically, } M_X: \{\underline{x} = \Phi_{(\varphi,X)}(\underline{\omega}) \,|\, \underline{\omega} \in M\} \text{ and } M_Y: \left\{\underline{y} = \Phi_{(\varphi,Y)}(\underline{\omega}) \,\middle|\, \underline{\omega} \in M\right\}.$$

Since $\exists\, \psi: M_X \rightarrowtail M_Y$, with ψ bijective (injective and surjective), this means

(1) $\forall \underline{x}, \underline{x}' \in M_X,\ \psi(\underline{x}) = \psi(\underline{x}') \Longrightarrow \underline{x} = \underline{x}'$, or equivalently $\forall \underline{x}, \underline{x}' \in M_X,\ \underline{x} \neq \underline{x}' \Longrightarrow \psi(\underline{x}) \neq \psi(\underline{x}')$, and

(2) $\forall \underline{y} \in M_Y, \exists!\ \underline{x} \in M_X, \psi(\underline{x}) = \underline{y}$

From (2), it is obvious that, since every point $\underline{y}$ in M_Y is determined via ψ from a unique point $\underline{x}$ from M_X, M_X causes M_Y in line with Definition 4.1. ■

The following lemma provides the necessary ingredients for us to logically deduce influence from shadow attractors on their respective time series.

Lemma 4.2. *Let M be an m-dimensional compact manifold and $X: M \to \mathbb{R}$, $Y: M \to \mathbb{R}$ be smooth observation functions. If M_X causes M_Y (through a bijective map ψ) and there exist $h: M_X \twoheadrightarrow X$ and $g: M_Y \twoheadrightarrow Y$, with h, g bijective, then X causes Y as well.*

Proof of Lemma 4.2: According to Lemma 1, since M_X causes M_Y, there exists a bijective map ψ such that $\forall \underline{y} \in M_Y, \exists!\ \underline{x} \in M_X, \psi(\underline{x}) = \underline{y}$.

Since h, g are bijective, $\forall \underline{x} \in M_X, \exists!\ x \in X, \mathrm{h}(\underline{x}) = x$, and $\forall \underline{y} \in M_Y, \exists!\ y \in Y, \mathrm{g}\left(\underline{y}\right) = y$.

Thus,

$\psi(\underline{x}) = \underline{y} \Leftrightarrow \psi(\,\mathrm{h}(x)) = \mathrm{g}(y) \overset{g^{-1}:bijective}{\Longleftrightarrow} (g^{-1} \circ \psi \circ h)(x) = (g^{-1} \circ g)(y) \Leftrightarrow y = (g^{-1} \circ \psi \circ h)(x)$.

By setting $(g^{-1} \circ \psi \circ h) = \rho, y = \rho(x), \forall y \in Y$.

From the last equation, it is obvious that, since every y in Y is determined via ρ from a unique point x from X, X causes Y, in line with Definition 4.1. ■

Remark: The composition of bijective functions is a bijection.

As articulated in Lemmas 4.1 and 4.2, strong influence from X to Y is established by X having an influence on all values of Y. This is achieved primarily through the bijective property firstly of $\psi: M_X \twoheadrightarrow M_Y$ and subsequently of $h: M_X \twoheadrightarrow X$ and $g: M_Y \twoheadrightarrow Y$. Bijection guarantees that, ultimately, at every time step of the dynamical evolution, Y is influenced by X, thus X strongly causes Y. The following two lemmas are similar to Lemmas 4.1 and 4.2, with the only difference being that we relax the requirement for a bijective mapping to an injective mappying, and thus the deduction is "weak" influence. This form of influence is what we expect to find in real data.

Lemma 4.3. *Let M be an m-dimensional compact manifold and $X: M \to \mathbb{R}$, $Y: M \to \mathbb{R}$ be smooth observation functions. Let $\varphi: M \rightarrowtail M$ be a smooth diffeomorphism. If there exists $\psi: M_X \rightarrowtail M_Y$ such that ψ is injective, then M_X weakly causes M_Y.*

Proof of Lemma 4.3: Since X, Y are smooth functions and φ is a smooth diffeomorphism, according to Takens' theorem, it is a generic property that the maps $\Phi_{(\varphi,X)}(\underline{\omega}): M \to \mathbb{R}^{2m+1}$ and $\Phi_{(\varphi,Y)}(\underline{\omega}): M \to \mathbb{R}^{2m+1}$, defined as

$$\Phi_{(\varphi,X)}(\underline{\omega}) = \langle X(\underline{\omega}), X(\varphi(\underline{\omega})), \dots, X(\varphi^{2m}(\underline{\omega})) \rangle$$

$$\Phi_{(\varphi,Y)}(\underline{\omega}) = \langle Y(\underline{\omega}), Y(\varphi(\underline{\omega})), \dots, Y(\varphi^{2m}(\underline{\omega})) \rangle$$

are embeddings (reconstructions) of the original manifold M and $\underline{\omega} \in M$.

Specifically, $M_X: \{\underline{x} = \Phi_{(\varphi,X)}(\underline{\omega}) | \underline{\omega} \in M\}$ and $M_Y: \left\{\underline{y} = \Phi_{(\varphi,Y)}(\underline{\omega}) \middle| \underline{\omega} \in M\right\}$.

Since $\exists\ \psi: M_X \rightarrowtail M_Y$, with ψ being injective, this means

(1) $\forall \underline{x}, \underline{x}' \in M_X,\ \psi(\underline{x}) = \psi(\underline{x}') \Longrightarrow \underline{x} = \underline{x}'$, or equivalently, $\forall \underline{x}, \underline{x}' \in M_X,\ \underline{x} \neq \underline{x}' \Longrightarrow \psi(\underline{x}) \neq \psi(\underline{x}')$, and

(2) $\forall \underline{y} \in M_Y,\ \exists_{\leq 1}\ \underline{x} \in M_X, \psi(\underline{x}) = \underline{y}$

From (2), it is obvious that every point $\underline{y}$ in M_Y is determined by at most one $\underline{x} \in M_X$ (it is not guaranteed, though, that every $\underline{y}$ is determined by some $\underline{x}$, thus the influence is "weak", unlike with a bijective ψ). Thus M_X causes M_Y, in line with Definition 4.1. ■

Lemma 4.4. *Let M be an m-dimensional compact manifold and $X: M \to \mathbb{R}$, $Y: M \to \mathbb{R}$ be smooth observation functions. If M_X weakly causes M_Y (through an injective map ψ) and there exist $h: M_X \rightarrowtail X$ and $g: M_Y \rightarrowtail Y$ with h, g being injective, then X weakly causes Y as well.*

Proof of Lemma 4.4: According to Lemma 4.3, since M_X causes M_Y, there exists an injective map ψ such that $\forall \underline{y} \in M_Y,\ \exists_{\leq 1}\ \underline{x} \in M_X, \psi(\underline{x}) = \underline{y}$.

Since h, g are injective, $\forall \underline{x} \in M_X,\ \exists_{\leq 1}\ x \in X, \mathrm{h}(\underline{x}) = x$, and $\forall \underline{y} \in M_Y,\ \exists_{\leq 1}\ y \in Y, \mathrm{g}\left(\underline{y}\right) = y$.

Thus,

$\psi(\underline{x}) = \underline{y} \Leftrightarrow \psi(\,\mathrm{h}(x)) = \mathrm{g}(y) \overset{g^{-1}:injective}{\Longleftrightarrow} (g^{-1} \circ \psi \circ h)(x) = (g^{-1} \circ g)(y) \Leftrightarrow y = (g^{-1} \circ \psi \circ h)(x)$.

By setting $(g^{-1} \circ \psi \circ h) = \rho$, $y = \rho(x)$, for some $y \in Y$.

From the last equation, it is obvious that some y in Y are determined via ρ from a unique point x from *X*. Thus, *X* causes *Y*, in line with Definition 4.1. ■

Remark: The composition of injective functions is an injection.

By relaxing the requirement from a bijection (see Lemmas 4.1 and 4.2) to an injection (see Lemmas 4.3 and 4.4), we expect that, for *X* to weakly cause *Y*, essentially, *ψ, h* and *g* have to be injective. Injection suggests that, at some time steps of the dynamical evolution, *Y* is influenced by *X*, thus *X* weakly causes *Y*, and the strength of influence is determined by the frequency of the mapping. Having established the prerequisite lemmas, we are now in a position to develop the main results that allow the nature of influence to be expressed and quantified. To that end, we need to use patterns from symbolic dynamics theory.

Now filtering the vectors $\underline{x}(t) \in M_X$ and $\underline{y}(t) \in M_Y$ through symbolic dynamics, we can extract their corresponding patterns $P_{\underline{x}(t)}$ and $P_{\underline{y}(t)}$ and we can distinguish three types of mappings (see Tables 4.7 and 4.8).

Definition 4.2.

(1) $\mathcal{P}: M_X \overset{+}{\rightarrow} M_Y$, $\overset{+}{\rightarrow}: \{\underline{x} \rightarrow \underline{y} | P_{\underline{x}} \overset{+}{\leftrightarrow} P_{\underline{x}}\}$, $\mathcal{P}$ corresponds to the same patterns (i.e. positive mapping)

(2) $\mathcal{N}: M_X \overset{-}{\rightarrow} M_Y$, $\overset{-}{\rightarrow}: \{\underline{x} \rightarrow \underline{y} | P_{\underline{x}} \overset{-}{\leftrightarrow} P_{\underline{x}}\}$, $\mathcal{N}$ corresponds to opposite patterns (i.e. negative mapping)

(3) $\mathcal{D}: M_X \overset{\star}{\rightarrow} M_Y$, $\overset{\star}{\rightarrow}: \{\underline{x} \rightarrow \underline{y} | P_{\underline{x}} \overset{\star}{\leftrightarrow} P_{\underline{x}}\}$, $\mathcal{D}$ corresponds to patterns that are neither the same nor opposite (i.e. dark mapping)

Definition 4.2 is used extensively in the formulation and proof of the following three important theorems.

Theorem 4.1. *Let M be an m-dimensional compact manifold and $X: M \to \mathbb{R}$, $Y: M \to \mathbb{R}$ be smooth observation functions. Let $\varphi: M \nrightarrow M$ be a smooth diffeomorphism. Let M_X and M_Y be the shadow attractors of X and Y respectively. If $\mathcal{P}: M_X \overset{+}{\to} M_Y$ such that $\mathcal{P}$ is bijective (or injective) and there exist $h: M_X \to X$ and $g: M_Y \to Y$, with h, g being bijective (or injective), then X exerts a positive influence on Y.*

Proof of Theorem 4.1: Since X, Y are smooth functions and φ is a smooth diffeomorphism, according to Takens' theorem it is a generic property that the maps $\Phi_{(\varphi,X)}(\underline{\omega}): M \to \mathbb{R}^{2m+1}$ and $\Phi_{(\varphi,Y)}(\underline{\omega}): M \to \mathbb{R}^{2m+1}$are embeddings (reconstructions) of the original manifold M and $\underline{\omega} \in M$. Since $\exists\, \mathcal{P}: M_X \overset{+}{\to} M_Y$, with $\mathcal{P}$ being bijective (injective), according to Lemma 4.1 (or similarly Lemma 4.3), M_X causes M_Y, and since $\exists h, g$ that are bijective (injective), i.e., $h: M_X \to X$ and $g: M_Y \to Y$, according to Lemma 4.2 (or similarly Lemma 4.4), X causes Y. Since $\mathcal{P}$ refers only to same-pattern couplings, according to Definition 1, X exerts a positive influence on Y. ■

Theorem 4.2. *Let M be an m-dimensional compact manifold and $X: M \to \mathbb{R}$, $Y: M \to \mathbb{R}$ be smooth observation functions. Let $\varphi: M \nrightarrow M$ be a smooth diffeomorphism. Let M_X and M_Y be the shadow attractors of X and Y respectively. If $\mathcal{N}: M_X \overset{-}{\to} M_Y$ is bijective (or injective) and there exist $h: M_X \to X$ and $g: M_Y \to Y$ that are bijective (or injective), then X exerts a negative influence on Y.*

Proof of Theorem 4.2: Since X, Y are smooth functions and φ is a smooth diffeomorphism, according to Takens' theorem, it is a generic property that the maps $\Phi_{(\varphi,X)}(\underline{\omega}): M \to \mathbb{R}^{2m+1}$ and $\Phi_{(\varphi,Y)}(\underline{\omega}): M \to \mathbb{R}^{2m+1}$are embeddings (reconstructions) of the original manifold M and $\underline{\omega} \in M$. Since $\exists\, \mathcal{N}: M_X \overset{-}{\to} M_Y$, with $\mathcal{N}$ being bijective (injective), according to Lemma 4.1 (or similarly Lemma 4.3), M_X causes M_Y, and since $\exists h, g$ that are bijective (injective), i.e., $h: M_X \to X$ and $g: M_Y \to Y$, according to Lemma 4.2 (or similarly Lemma 4.4), X causes Y. Since $\mathcal{N}$ refers only to opposite-pattern couplings, according to Definition 4.1, X exerts a negative influence on Y. ■

Theorem 4.3. *Let M be an m-dimensional compact manifold and $X: M \to \mathbb{R}$, $Y: M \to \mathbb{R}$ be smooth observation functions. Let $\varphi: M \nrightarrow M$ be a smooth diffeomorphism. Let M_X and M_Y be*

the shadow attractors of X and Y respectively. If $\mathcal{D}: M_X \overset{\star}{\to} M_Y$ *such that* $\mathcal{D}$ *is bijective (or injective) and there exist* $h: M_X \to X$ *and* $g: M_Y \to Y$*, with* h, g *being bijective (or injective), then X exerts a dark influence on* Y.

Proof of Theorem 4.3: Since X, Y are smooth functions and φ is a smooth diffeomorphism, according to Takens' theorem, it is a generic property that the maps $\Phi_{(\varphi,X)}(\underline{\omega}): M \to \mathbb{R}^{2m+1}$ and $\Phi_{(\varphi,Y)}(\underline{\omega}): M \to \mathbb{R}^{2m+1}$are embeddings (reconstructions) of the original manifold M and $\underline{\omega} \in M$. Since $\exists\, \mathcal{D}: M_X \overset{\star}{\to} M_Y$, with $\mathcal{D}$ being bijective (injective), according to Lemma 4.1 (or similarly Lemma 4.3), M_X causes M_Y, and since $\exists h, g$ that are bijective (injective), i.e., $h: M_X \to X$ and $g: M_Y \to Y$, according to Lemma 4.2 (or similarly Lemma 4.4), X causes Y. Since $\mathcal{D}$ refers only to pattern couplings which are neither the same nor opposite, according to Definition 4.1, X exerts a dark influence on Y. ■

4.11 Signature Calculation and Pattern Causality Matrix

Let us have four patterns:

- $s_1 = \nearrow\nearrow = (0.32, 0.45)$, with corresponding weight, $w_1 = 0.91$.
- $s_2 = \searrow\nearrow = (-0.11, 0.51)$, with corresponding weight, $w_2 = 0.54$.
- $s_3 = \nearrow\nearrow = (0.13, 0.19)$, with corresponding weight, $w_3 = 0.82$.
- $s_4 = \nearrow\searrow = (0.05, -0.08)$, with corresponding weight, $w_4 = 0.69$.

The weighted average in our example is

$S = \sum_{i=1}^{4} w_i\, s_i = 0.91 * (0.32, 0.45) + 0.54 * (-0.11, 0.51) + 0.82 * (0.13, 0.19) + 0.69 * (0.05, -0.08) = (0.3729, 0.7855) = (\nearrow\nearrow)$

Thus, the emergent weighted average pattern is the signature of S: $\boldsymbol{P}$=$signature(S)$=$\nearrow\nearrow$.

Table 4.6: PC (from X to Y) pattern to pattern matrix for E = 2. Each cell is filled with the accuracy from step 6 of our algorithm. Thus, each cell takes values from 0 to 1. Blue cells denote positive influence. Red cells denote negative influence. Purple cells denote dark influence.

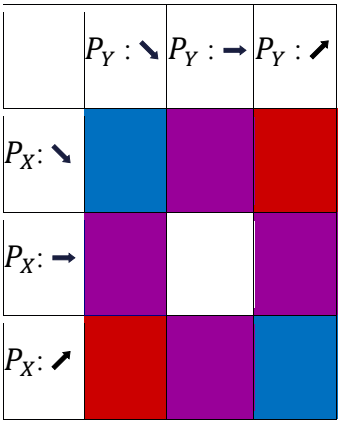
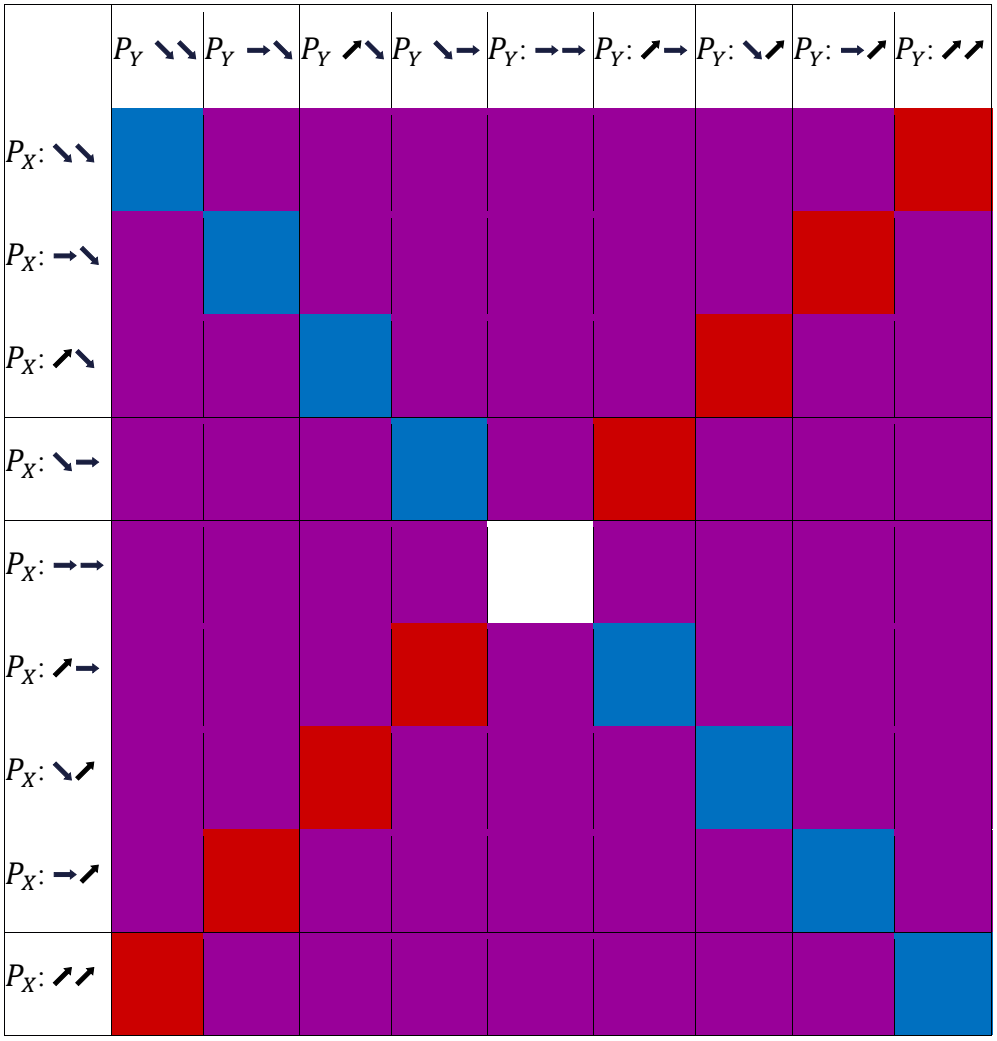


Table 4.7: PC (from X to Y) pattern to pattern matrix for E = 3. Each cell is filled with the accuracy from step 6 of our algorithm. Thus, each cell takes values from 0 to 1. Blue cells denote positive influence. Red cells denote negative influence. Purple cells denote dark influence.



5 General Conclusions and Future Research Plans

The theme of this research is centered on financial networks and their emergent intricacies. We are seeking real world meaning imprinted in persistent but hidden network structures. Our theoretical discourse suggests that despite the apparent randomness of single assets, when studied collectively as a network, they exhibit persistent patterns that signify crucial messages for scientists and policymakers alike.

As a first step, we collected eight causality methods from a wide literature spectrum. After a thorough analysis using these methods in a custom portfolio of assets, we identified the existence of crucial interdependencies between certain financial assets. Our findings suggest that two specific nonlinear causality tools produce distinctive features before and during the financial crisis of 2007-9. These results challenge the efficient market hypothesis and open new horizons for further investigation and possible arbitrage opportunities.

Specifically, such arbitrage opportunities would entail the causality analysis of collectives of asset prices, in effect candidate assets for portfolio design. Ideally a dashboard of the causal structure of financial assets would dynamically inform the investor for persistent pairs ripe for pairs trading, or for the timely situation of buying or selling a single asset when this asset is strongly influenced by many other assets, thus rendering its imminent price change quite predictable. The common denominator of any strategy stemming from causality analysis is that one takes a position on a given asset, based not only on information about the asset's past performance but also on the causal dependencies that this asset is entangled into. Causality can be a tool through which the prudent investor can envelop the asset's price fluctuations at hand with interdependencies that most investors are unaware.

Notwithstanding the insights from the established causality methods, the hidden nature of causality remains a puzzling yet critical notion for effective decision-making. Financial markets are characterized by fluctuating interdependencies that seldom give rise to emergent phenomena such as bubbles or crashes. Thus, addressing the complex needs of nonlinear time-series properties, we design a novel causality framework based on symbolic dynamics that probes beneath the surface of abstract causality and unveils the nature of causal interactions. We named our framework "pattern causality".

Pattern causality is a method which unveils the structure of complex systems through time series data. Thus, taking a pair of time series and testing it for causality, we check whether,

how much, and in which direction *X* causes *Y*. In this regard, first we reconstruct the shadow attractors, i.e., their time-delayed representations on at least a 2-dimensional space. Finally, we test *X*'s ability to predict *Y*'s values. The better the prediction accuracy is the stronger the causality from *X* to *Y*.

This novel algorithm allows distinction between positive and negative interdependencies as well as a hybrid form that we refer to as "dark causality". Initially our framework is validated by models of *a priori* defined causal interdependencies. Then we test our method on asset pairs and on a network of sovereign credit default swaps (CDS). Our findings suggest that "dark causality" dominates the sovereign CDS network, indicating interdependencies that require caution from an investor's perspective.

Pointedly, dark causality in a pair of assets is the quantified multifarious influence that one asset has on the other. When dark causality is found we know that one asset influences another, however the direction of influence is mixed. Mixed means that causality is neither positive nor negative (the sign of positive or negative follows, very liberally, the correlation rationale). However, since dark causality, is still a form of influence though quite nonlinear, its implications has not been revealed yet in the literature. That is why, when investors encounter dark causal relationships in their portfolios, they should be very cautious, because strategies and tactics to deal with dark causality have yet to be developed.

The implications of dark causality on the decision process of regulators and policy makers are even more momentous. Given that interdependencies among assets can lead to bubbles and crashes, the far-reaching existence of dark causality in the global financial network foreshadows a deep uncertainty as to what to expect for the imminent state of the markets. This fact raises the necessity to equip decision making frameworks with a novel arsenal ready to capture early warning signals of dark causality but most importantly to provide ways to anticipate and regulate the financial system against incoming crises.

Further contemplating upon the operational laws and concepts from complex systems, we composed a second algorithm out of the pattern causality framework with the purpose of capturing important aspects and interactions beyond stock markets. In an abstract complex network, it remains enigmatic and challenging, yet inspiring, to predict the actual interdependencies that comprise the structure of such systems, be it financial markets, ecosystems, or even the human brain. Particularly considering that the vital interdependencies underlying disparate real-world phenomena might be persistently hidden, the task of creating

one algorithm to tackle them all seems daunting. Yet, our second algorithm is excellent at detecting the latent and elusive structures of complex systems. Our treatment utilizes short-term predictions from information embedded in reconstructed state space. Using a broad class of real-world applications, we are able to demonstrate our method's power to reconstruct the backbone of complex systems and simultaneously highlight their most fundamental operations. This last algorithm can serve as a tool for decision-makers and policymakers and the demonstrated effectiveness establishes its excellent potential for capturing hidden interactions in a much broader area of applications.

The detection of dark causality is something new, yet we believe it will spark the beginning of a new field in the area of complex networks. This field will be characterized by quantitative methods for a deeper and more robust understanding of variable interdependencies beyond simple statistics. Our framework and its two algorithms are the prelude in this direction.

We envision three main paths for future research based upon our foundations. The first and more straightforward is that of applying our framework in even more challenging settings in the field of finance, possibly in the direction of stock market forecasting. The second and more general path entails applying our framework in disciplines beyond finance such as ecology, neuroscience, and even sociology. Finally, the last and more challenging path is towards the expansion and enhancement of our framework to accommodate new algorithms with the purpose of identifying hidden structures in networks from diverse fields.

Especially in the field of finance, the detection of dark causality has a direct impact in the portfolio design. Using the pattern causality framework, we can create a financial theory that will indicate optimal portfolios based on the spectrum of positive, negative and dark linkages among financial assets. This finer detail of interdependencies will allow for the customization of portfolios based on the ratio of causality types, according to the expectations that such causality blends will have. For example, an investor who wants to speculate all-or-nothing would build a portfolio that maximizes positive causality. On the other hand, an investor who desires diversification strategies or pairs trading tactics would aim for negative causality maximization. Finally, a portfolio which maximizes dark causality has yet to be discovered implications, and this in itself should be a separate field of study.

Bibliography

Abhyankar, A. (1998). Linear and nonlinear Granger causality: Evidence from the UK stock index futures market. The Journal of Futures Markets (1986-1998), 18(5), 519. DOI: https://doi.org/10.1002/(SICI)1096-9934(199808)18:5<519::AID-FUT2>3.0.CO;2-U

Abhyankar, A., Copeland, L. S., & Wong, W. (1997). Uncovering nonlinear structure in real-time stock-market indexes: the S&P 500, the DAX, the Nikkei 225, and the FTSE-100. Journal of Business & Economic Statistics, 15(1), 1-14. DOI: https://doi.org/10.1080/07350015.1997.10524681

Acharya, V., Philippon, T., Richardson, M., & Roubini, N. (2009). The financial crisis of 2007-2009: Causes and remedies. Financial markets, institutions & instruments, 18(2), 89-137. DOI: https://doi.org/10.1111/j.1468-0416.2009.00147_2.x

Alexander, C. (2001). Market models: A guide to financial data analysis. John Wiley & Sons.

Alligood, K. T., Sauer, T. D., Yorke, J. A., & Chillingworth, D. (1998). Chaos: An Introduction to Dynamical Systems. SIAM Review, 40(3), 732-732. http://weber.itn.liu.se/~krzma/DS2018/Alligood%20Sauer%20and%20Yorke%20-%20%20Chaos%20An%20Introduction%20to%20Dynamical%20Systems.pdf

Allington, G. R., Koons, D. N., Morgan Ernest, S. K., Schutzenhofer, M. R., & Valone, T. J. (2013). Niche opportunities and invasion dynamics in a desert annual community. Ecology letters, 16(2), 158-166. DOI: https://doi.org/10.1111/ele.12023

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders. BMC Med, 17, 133-137. DOI: https://doi.org/10.4103/0019-5545.117131

Apergis, N., & Payne, J. E. (2014). Resurrecting the size effect: Evidence from a panel nonlinear cointegration model for the G7 stock markets. Review of Financial economics, 23(1), 46-53. DOI: https://doi.org/10.1016/j.rfe.2013.08.003

Apostle, H. G. (Ed.). (1966). Metaphysics. Indiana University Press. DOI: https://doi.org/10.2307/2218565

Arnold, G. (2012). Financial Times Guide to the Financial Markets. Pearson Education.

Arthur, W. B. (1990). Positive feedbacks in the economy. Scientific american, 262(2), 92-99. DOI: http://blackout.gmu.edu/archive/pdf/arthur_90.pdf

Aste, T., & Di Matteo, T. (2006). Dynamical networks from correlations. Physica A: Statistical Mechanics and its Applications, 370(1), 156-161. DOI: https://doi.org/10.1016/j.physa.2006.04.019

Atchison, M. D., Butler, K. C., & Simonds, R. R. (1987). Nonsynchronous security trading and market index autocorrelation. The Journal of Finance, 42(1), 111-118. DOI: https://doi.org/10.1111/j.1540-6261.1987.tb02553.x

Athanasenas, A., Katrakilidis, C., & Trachanas, E. (2014). Government spending and revenues in the Greek economy: evidence from nonlinear cointegration. Empirica, 41(2), 365-376. DOI: http://doi.org/b7mx

Babihuga, R., & Spaltro, M. (2014). Bank funding costs for international banks (No. 14-71). International Monetary Fund. DOI: https://www.imf.org/external/pubs/ft/wp/2014/wp1471.pdf

Baek, S. K., Jung, W. S., Kwon, O., & Moon, H. T. (2005). Transfer entropy analysis of the stock market. arXiv preprint physics/0509014. DOI: https://arxiv.org/pdf/physics/0509014.pdf

Balke, N. S., & Fomby, T. B. (1997). Threshold cointegration. International economic review, 627-645. DOI: https://doi.org/10.2307/2527284

Bandt, C., & Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. Physical review letters, 88(17), 174102. DOI: https://doi.org/10.1103/PhysRevLett.88.174102

Bandt, C., & Pompe, B. (1993). The entropy profile—A function describing statistical dependences. Journal of statistical physics, 70(3-4), 967-983. DOI: https://doi.org/10.1007/BF0105360

Barnes, J. A., & Harary, F. (1983). Graph theory in network analysis. DOI: https://doi.org/10.1016/0378-8733(83)90026-6

Barrat, A., Barthelemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. Proceedings of the national academy of sciences, 101(11), 3747-3752. DOI: https://doi.org/10.1073/pnas.0400087101

Bavelas, A. (1950). Communication patterns in task-oriented groups. The Journal of the Acoustical Society of America, 22(6), 725-730. DOI: https://doi.org/10.1121/1.1906679

Begleiter, H., Porjesz, B., & Wang, W. (1993). A neurophysiologic correlate of visual short-term memory in humans. Electroencephalography and clinical neurophysiology, 87(1), 46-53. DOI: https://doi.org/10.1016/0013-4694(93)90173-S

Bellman, R. E., & Zadeh, L. A. (1970). Decision-making in a fuzzy environment. Management science, 17(4), B-141. DOI: https://doi.org/10.1287/mnsc.17.4.B141

Benhmad, F. (2012). Modeling nonlinear Granger causality between the oil price and US dollar: A wavelet-based approach. Economic Modelling, 29(4), 1505-1514. DOI: https://doi.org/10.1016/j.econmod.2012.01.003

Billio, M., Getmansky, M., Lo, A. W., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. Journal of financial economics, 104(3), 535-559. DOI: https://doi.org/10.1016/j.jfineco.2011.12.010

Birch, J., Pantelous, A. A., & Zuev, K. (2015). The maximum number of 3-and 4-cliques within a planar maximally filtered graph. Physica A: Statistical Mechanics and its Applications, 417, 221-229. DOI: https://doi.org/10.1016/j.physa.2014.09.011

Boginski, V., Butenko, S., & Pardalos, P. M. (2005). Statistical analysis of financial networks. Computational statistics & data analysis, 48(2), 431-443. DOI: https://doi.org/10.1016/j.csda.2004.02.004

Bradshaw, G. W., & Orden, D. (1990). Granger causality from the exchange rate to agricultural prices and export sales. Western Journal of Agricultural Economics, 100-110. DOI: http://bit.ly/2rDxWIs.

Breese, G. R., Sinha, R., & Heilig, M. (2011). Chronic alcohol neuroadaptation and stress contribute to susceptibility for alcohol craving and relapse. Pharmacology & therapeutics, 129(2), 149-171. DOI: https://doi.org/10.1016/j.pharmthera.2010.09.007

Buch, C. M., Koetter, M., & Ohls, J. (2016). Banks and sovereign risk: A granular view. Journal of Financial Stability, 25, 1-15. DOI: https://doi.org/10.1016/j.jfs.2016.05.002

Casdagli, M. (1989). Nonlinear prediction of chaotic time series. Physica D: Nonlinear Phenomena, 35(3), 335-356. DOI: https://doi.org/10.1016/0167-2789(89)90074-2

Cellucci, C. J., Albano, A. M., & Rapp, P. E. (2003). Comparative study of embedding methods. Physical Review E, 67(6), 066210. DOI: https://doi.org/10.1103/PhysRevE.67.066210

Cerchi, M., & Havenner, A. (1988). Cointegration and stock prices: The random walk on Wall Street revisited. Journal of Economic Dynamics and Control, 12(2-3), 333-346. DOI: https://doi.org/10.1016/0165-1889(88)90044-9

Chapin III, F. S., Matson, P. A., & Vitousek, P. (2011). Principles of terrestrial ecosystem ecology. Springer Science & Business Media.

Chow, G., Jacquier, E., Kritzman, M., & Lowry, K. (1999). Optimal portfolios in good times and bad. Financial Analysts Journal, 55(3), 65-73. DOI: https://doi.org/10.2469/faj.v55.n3.2273

Chowdhury, A. R. (1991). Futures market efficiency: evidence from cointegration tests. Journal of Futures markets, 11(5), 577-589. DOI: https://doi.org/10.1002/fut.3990110506

Christensen, E. M., Harris, D. J., & Ernest, S. M. (2018). Long-term community change through multiple rapid transitions in a desert rodent community. Ecology, 99(7), 1523-1529. DOI: https://doi.org/10.1002/ecy.2373

Chu, X., Wu, C., & Qiu, J. (2016). A nonlinear Granger causality test between stock returns and investor sentiment for Chinese stock market: a wavelet-based approach. *Applied Economics*, *48*(21), 1915-1924. DOI: https://doi.org/10.1080/00036846.2015.1109048

Coleman, T. F matrices and graph coloring blems. SIAM journal on Numerical Analysis, 20(1), 187-209. DOI: https://doi.org/10.1137/0720013

Crawford G., & Sen, B. (1996). Derivatives for decision makers: Strategic management issues (Vol. 1). John Wiley & Sons.

Crespi, B. J. (2004). Vicious circles: positive feedback in major evolutionary and ecological transitions. Trends in Ecology & Evolution, 19(12), 627-633. DOI: https://doi.org/10.1016/j.tree.2004.10.001

Crutchfield, J. P., & Young, K. (1989). Inferring statistical complexity. Physical Review Letters, 63(2), 105. DOI: https://doi.org/10.1103/PhysRevLett.63.105

Curme, C., Tumminello, M., Mantegna, R. N., Stanley, H. E., & Kenett, D. Y. (2015). Emergence of statistically validated financial intraday lead-lag relationships. Quantitative Finance, 15(8), 1375-1386. DOI: https://doi.org/10.1080/14697688.2015.1032545

Da, Z., Engelberg, J. and Gao, P., 2011. In search of attention. The Journal of Finance, 66(5), 1461-1499. DOI: https://doi.org/10.1111/j.1540-6261.2011.01679.x

Da Silva, F. L., Pijn, J. P., & Boeijinga, P. (1989). Interdependence of EEG signals: linear vs. nonlinear associations and the significance of time delays and phase shifts. *Brain topography*, *2*(1-2), 9-18. DOI: http://doi.org/d6gdth

Deyle, E. R., & Sugihara, G. (2011). Generalized theorems for nonlinear state space reconstruction. PLoS One, 6(3), e18295. DOI: https://doi.org/10.1371/journal.pone.0018295

Deyle, E. R., May, R. M., Munch, S. B., & Sugihara, G. (2016). Tracking and forecasting ecosystem interactions in real time. Proceedings of the Royal Society B: Biological Sciences, 283(1822), 20152258. DOI: https://doi.org/10.1098/rspb.2015.2258

Di Matteo, T., Pozzi, F., & Aste, T. (2010). The use of dynamical networks to detect the hierarchical organization of financial market sectors. The European Physical Journal B, 73(1), 3-11. DOI: http://doi.org/d2rh22

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica: journal of the Econometric Society, 1057-1072. DOI: http://doi.org/fbr6s5

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American statistical association, 74(366a), 427-431. DOI: https://doi.org/10.1080/01621459.1979.10482531

Dionisio, A. T., Menezes, R., Mendes, D., & Vidigal da Silva, J. (2007). Nonlinear Dynamics within Macroeconomic Factors and Stock Market in Portugal, 1993-2003. Applied Econometrics and International Development, 7(2). DOI: http://bit.ly/2rZcQRq.

Dutta, A. (2001). Telecommunications and economic activity: An analysis of Granger causality. Journal of Management Information Systems, 17(4), 71-95. DOI: https://doi.org/10.1080/07421222.2001.11045658

Eagle, N., Macy, M., & Claxton, R. (2010). Network diversity and economic development. Science, 328(5981), 1029-1031. DOI: https://doi.org/10.1126/science.1186605

Edelstein-Keshet, L. (2004). Mathematical models in biology, SIAM Classics Appl. Math, 46, 586. DOI: https://trove.nla.gov.au/work/12921600

Engelberg, J. O. S. E. P. H., & Gao, P. (2011). In search of attention. The Journal of Finance, 66(5), 1461-1499. DOI: https://doi.org/10.1111/j.1540-6261.2011.01679.x

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. Econometrica: journal of the Econometric Society, 251-276. DOI: https://doi.org/10.2307/1913236

Eom, C., Jung, W. S., Choi, S., Oh, G., & Kim, S. (2008). Effects of time dependency and efficiency on information flow in financial markets. Physica A: Statistical Mechanics

and its Applications, 387(21), 5219-5224. DOI: https://doi.org/10.1016/j.physa.2008.05.054

Erkens, D. H., Hung, M., & Matos, P. (2012). Corporate governance in the 2007–2008 financial crisis: Evidence from financial institutions worldwide. Journal of corporate finance, 18(2), 389-411. DOI: https://doi.org/10.1016/j.jcorpfin.2012.01.005

Morgan Ernest, S. K., Yenni, G. M., Allington, G., Christensen, E. M., Geluso, K., Goheen, J. R., ... & Valone, T. J. (2016). Long-term monitoring and experimental manipulation of a Chihuahuan desert ecosystem near Portal, Arizona (1977–2013). Ecology, 97(4), 1082-1082. DOI: https://doi.org/10.1890/15-2115.1

Escanciano, J. C., & Escribano, A. (2009). Econometrics: Non-linear cointegration. Encyclopedia of Complexity and Systems Science, 2757-2769. DOI: https://doi.org/10.1007/978-0-387-30440-3_166

Escribano, A., & Mira, S. (2002). Nonlinear error correction models. Journal of Time Series Analysis, 23(5), 509-522. DOI: https://doi.org/10.1111/1467-9892.00276

Farmer, J. D., & Sidorowich, J. J. (1987). Predicting chaotic time series. Physical review letters, 59(8), 845. DOI: https://doi.org/10.1103/PhysRevLett.59.845

Fiedor, P. (2014). Networks in financial markets based on the mutual information rate. Physical Review E, 89(5), 052801. DOI: https://doi.org/10.1103/PhysRevE.89.052801

Fiedor, P. (2015). Granger-causal nonlinear financial networks. Journal of Network Theory in Finance, 1(2), 53-82. DOI: https://doi.org/10.1103/PhysRevE.89.052801

Fiedor, P. (2014). Information-theoretic approach to lead-lag effect on financial markets. The European Physical Journal B, 87(8), 168. DOI: http://doi.org/b7m6

Foresti, P. (2006). Testing for Granger causality between stock prices and economic growth. DOI: http://bit.ly/2cjnomb

Frank, M., & Stengos, T. (1989). Measuring the strangeness of gold and silver rates of return. The Review of Economic Studies, 56(4), 553-567. DOI: https://doi.org/10.2307/2297500

Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. Physical review A, 33(2), 1134. DOI: https://doi.org/10.1103/PhysRevA.33.1134

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. Sociometry, 35-41. DOI: http://dx.doi.org/10.2307/3033543

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. Social networks, 1(3), 215-239. DOI: http://dx.doi.org/ 10.1.1.227.9549

Froot, K. A., & Obstfeld, M. (1991). Exchange-rate dynamics under stochastic regime shifts: A unified approach. Journal of International Economics, 31(3-4), 203-229. DOI: https://doi.org/10.1016/0022-1996(91)90036-6

Frydman, C. and Camerer, C.F., (2016). The psychology and neuroscience of financial decision making. Trends in cognitive sciences, 20(9), pp.661-675. https://doi.org/10.1016/j.tics.2016.07.003

Gabaix, X., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003). A theory of power-law distributions in financial market fluctuations. Nature, 423(6937), 267. DOI: https://doi.org/10.1038/nature01624

Gai, P., & Kapadia, S. (2010). Contagion in financial networks. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 466(2120), 2401-2423. DOI: https://doi.org/10.1098/rspa.2009.0410

Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. The Review of Financial Studies, 19(3), 797-827. DOI: https://doi.org/10.1093/rfs/hhj020

Gower, J. C., & Ross, G. J. (1969). Minimum spanning trees and single linkage cluster analysis. Journal of the Royal Statistical Society: Series C (Applied Statistics), 18(1), 54-64. DOI: https://doi.org/10.2307/2346439

Grammatikos, T., & Vermeulen, R. (2012). Transmission of the financial and sovereign debt crises to the EMU: Stock prices, CDS spreads and exchange rates. Journal of International Money and Finance, 31(3), 517-533. DOI: https://doi.org/10.1016/j.jimonfin.2011.10.004

Granger, C. W. J. (2001, January). Investigating causal relations by econometric models and cross-spectral methods. In Essays in econometrics (pp. 31-47). Harvard University Press. DOI: http://dx.doi.org/10.2307/1912791

Granger, C. W., & Hallman, J. (1991). Long memory series with attractors. Oxford Bulletin of Economics and Statistics, 53(1), 11-26. DOI: https://doi.org/10.1111/j.1468-0084.1991.mp53001002.x

Granger, C. W. (1981). Some properties of time series data and their use in econometric model specification. Journal of econometrics, 16(1), 121-130. DOI: https://doi.org/10.1016/0304-4076(81)90079-8

Granger, C., & Lin, J. L. (1994). Using the mutual information coefficient to identify lags in nonlinear models. Journal of time series analysis, 15(4), 371-384. DOI: https://doi.org/10.1111/j.1467-9892.1994.tb00200.x

Granger, C. W., & Weiss, A. A. (2001, January). Time series analysis of error-correction models. In Essays in econometrics (pp. 129-144). Harvard University Press. DOI: https://doi.org/10.1016/B978-0-12-398750-1.50018-8

Gündüz, Y., & Kaya, O. (2014). Impacts of the financial crisis on Eurozone sovereign CDS spreads. Journal of International Money and Finance, 49, 425-442. DOI: https://doi.org/10.1016/j.jimonfin.2014.03.013

Hakimi, S. L. (1963). On realizability of a set of integers as degrees of the vertices of a linear graph II. Uniqueness. Journal of the Society for Industrial and Applied Mathematics, 11(1), 135-147. DOI: https://doi.org/10.1137/0111010

Hakkio, C. S., & Rush, M. (1989). Market efficiency and cointegration: an application to the sterling and deutschemark exchange markets. Journal of international money and finance, 8(1), 75-88. DOI: https://doi.org/10.1016/0261-5606(89)90015-6

Hall, A. D., Anderson, H. M., & Granger, C. W. (1992). A cointegration analysis of treasury bill yields. The review of Economics and Statistics, 116-126. DOI: https://doi.org10.2307/2109549

Hall, J. E. (2015). Guyton and Hall textbook of medical physiology e-Book. Elsevier Health Sciences.

Hamilton, E., Cairns, H., & Cooper, L. (1961). The collected dialogues of Plato. Princeton University Press. DOI: https://muse.jhu.edu/book/46363

Hamilton, J. D. (1994). Time series analysis (Vol. 2, pp. 690-696). Princeton, NJ: Princeton university press. DOI: https://doc1.bibliothek.li/aca/FLMF037168.pdf

Hawawini, G. (1980). The intertemporal cross-price behavior of common stocks: Evidence and impications. DOI: https://mpra.ub.uni-muenchen.de/44896/

Hennig, B. (2009). The four causes. The journal of philosophy, 106(3), 137-160. DOI: http://dx.doi.org/10.5840/jphil200910634

Hertz, S., Porjesz, B., Begleiter, H., & Chorlian, D. (1994). Event-related potentials to faces: the effects of priming and recognition. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 92(4), 342-351. DOI: https://doi.org/10.1016/0168-5597(94)90102-3

Hiemstra, C., & Jones, J. D. (1994). Testing for linear and nonlinear Granger causality in the stock price-volume relation. The Journal of Finance, 49(5), 1639-1664. DOI: https://doi.org/10.1111/j.1540-6261.1994.tb04776.x

Hirsch, J. A., & Bishop, B. (1981). Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate. American Journal of Physiology-Heart and Circulatory

Physiology, 241(4), H620-H629. DOI: https://doi.org/10.1152/ajpheart.1981.241.4.H620

Hopcroft, J., & Tarjan, R. (1973). Algorithm 447: efficient algorithms for graph manipulation. Communications of the ACM, 16(6), 372-378. DOI: http://dx.doi.org/10.1145/362248.362272

Hsieh, D. A. (1989). Testing for nonlinear dependence in daily foreign exchange rates. Journal of Business, 62(3). DOI: http://www.finance.martinsewell.com/stylized-facts/nonlinearity/Hsieh1989.pdf

Hu, T. C. (1961). Letter to the editor—the maximum capacity route problem. Operations Research, 9(6), 898-900. DOI: https://doi.org/10.1287/opre.9.6.898

Huth, N., & Abergel, F. (2014). High frequency lead/lag relationships—empirical facts. Journal of Empirical Finance, 26, 41-58. DOI: https://doi.org/10.1016/j.jempfin.2014.01.003

Ignace, D. D., & Chesson, P. (2014). Removing an invader: evidence for forces reassembling a Chihuahuan Desert ecosystem. Ecology, 95(11), 3203-3212. DOI: https://doi.org/10.1890/14-0456.1

Iori, G., & Precup, O. V. (2007). Weighted network analysis of high-frequency cross-correlation measures. Physical Review E, 75(3), 036110. DOI: https://doi.org/10.1103/PhysRevE.75.036110

Iori, G., Mantegna, R. N., Marotta, L., Micciche, S., Porter, J., & Tumminello, M. (2015). Networked relationships in the e-MID Interbank market: A trading model with memory. Journal of Economic Dynamics and Control, 50, 98-116. DOI: https://doi.org/10.1016/j.jedc.2014.08.016

Iori, G., Reno, R., De Masi, G., & Caldarelli, G. (2007). Trading strategies in the Italian interbank market. Physica A: Statistical Mechanics and its Applications, 376, 467-479. DOI: https://doi.org/10.1016/j.physa.2006.10.053

Jacobson, L. D., & MacCall, A. D. (1995). Stock-recruitment models for Pacific sardine (Sardinops sagax). Canadian Journal of Fisheries and Aquatic Sciences, 52(3), 566-577. DOI: https://doi.org/10.1139/f95-057

Jiang, H. and Verardo, M., 2018. Does herding behavior reveal skill? An analysis of mutual fund performance. The Journal of Finance, 73(5), pp.2229-2269. https://doi.org/10.1111/jofi.12699

Josic, K. (2000). Synchronization of chaotic systems and invariant manifolds. Nonlinearity, 13(4), 1321. DOI: http://dx.doi.org/10.1103/PhysRevLett.88.174102

Junior, L., Mullokandov, A., & Kenett, D. (2015). Dependency relations among international stock market indices. *Journal of Risk and Financial Management*, *8*(2), 227-265. DOI: https://doi.org/10.3390/jrfm8020227

Kantz, H., & Schreiber, T. (2004). *Nonlinear time series analysis* (Vol. 7). Cambridge university press.

Kenett, D. Y., Tumminello, M., Madi, A., Gur-Gershgoren, G., Mantegna, R. N., & Ben-Jacob, E. (2010). Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market. *PloS one*, *5*(12), e15032. DOI: https://doi.org/10.1371/journal.pone.0015032

Kennel, M. B., Brown, R., & Abarbanel, H. D. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical review A*, *45*(6), 3403. DOI: https://doi.org/10.1103/PhysRevA.45.3403

Kim, J., Kim, G., An, S., Kwon, Y. K., & Yoon, S. (2013). Entropy-based analysis and bioinformatics-inspired integration of global economic information transfer. *PloS one*, *8*(1), e51986. DOI: https://doi.org/10.1371/journal.pone.0051986

Krakovská, A., Mezeiová, K., & Budáčová, H. (2015). Use of false nearest neighbours for selecting variables and embedding parameters for state space reconstruction. *Journal of Complex Systems*, *2015*. DOI: http://dx.doi.org/10.1155/2015/932750

Kroner, N. (2011). *A blueprint for better banking: Svenska Handelsbanken and a proven model for more stable and profitable banking*. Harriman House Limited.

Kullmann, L., Kertész, J., & Kaski, K. (2002). Time-dependent cross-correlations between different stock returns: A directed network of influence. *Physical Review E*, *66*(2), 026125. DOI: https://doi.org/10.1103/PhysRevE.66.026125

Kwan, S. H. (1996). Firm-specific information and the correlation between individual stocks and bonds. *Journal of financial economics*, *40*(1), 63-80. DOI: https://doi.org/10.1016/0304-405X(95)00836-4

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic

time series have a unit root?. *Journal of econometrics*, *54*(1-3), 159-178. DOI: https://doi.org/10.1016/0304-4076(92)90104-Y

Kwon, O., & Yang, J. S. (2008). Information flow between stock indices. *EPL (Europhysics Letters)*, *82*(6), 68003. DOI: http://dx.doi.org/10.1209/0295-5075/82/68003

Li, X. (2002). International stock market integration: evidence from nonlinear cointegration analysis. *Available at SSRN 961099*. DOI: http://dx.doi.org/10.2139/ssrn.961099

Liebert, W., & Schuster, H. G. (1989). Proper choice of the time delay for the analysis of chaotic time series. *Physics Letters A*, *142*(2-3), 107-111. DOI: https://doi.org/10.1016/0375-9601(89)90169-2

Liu, X., Song, H., & Romilly, P. (1997). Are Chinese stock markets efficient? A cointegration and causality analysis. *Applied Economics Letters*, *4*(8), 511-515. DOI: https://doi.org/10.1080/758536636

Lo, A. W., & MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction?. *The review of financial studies*, *3*(2), 175-205. DOI: https://doi.org/10.1093/rfs/3.2.175

Luce, R. D., & Perry, A. D. (1949). A method of matrix analysis of group structure. *Psychometrika*, *14*(2), 95-116. DOI: https://doi.org/10.1007/BF02289146

Luckinbill, L. S. (1973). Coexistence in laboratory populations of Paramecium aurelia and its predator Didinium nasutum. *Ecology*, *54*(6), 1320-1327. DOI: https://doi.org/10.2307/1934194

Ma, Y., & Kanas, A. (2004). Intrinsic bubbles revisited: Evidence from nonlinear cointegration and forecasting. *Journal of Forecasting*, *23*(4), 237-250. DOI: https://doi.org/10.1002/for.909

Maasoumi, E., & Wang, L. (2007). Economic reform, growth and convergence in China. *Econometrics Journal*, *10*(1), 1-25. DOI: https://doi.org/10.1111/j.1368-423X.2008.00233.x.

MacKay, R. S. (1991). Chaotic Evolution and Strange Attractors: The Statistical Analysis of Time Series for Deterministic Nonlinear Systems (David Ruelle). *SIAM Review*, *33*(2), 328. DOI: https://doi.org/10.1137/1033084

Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, *11*(1), 193-197. DOI: http://dx.doi.org/10.1007/s100510050929

Markowitz, H. (1952). Portfolio selection. *The journal of finance*, *7*(1), 77-91. DOI: https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

McClatchie, S., Goericke, R., Auad, G., & Hill, K. (2010). Re-assessment of the stock–recruit and temperature–recruit relationships for Pacific sardine (Sardinops sagax). *Canadian Journal of Fisheries and Aquatic Sciences*, *67*(11), 1782-1790. DOI: https://doi.org/10.1139/F10-101

Menezes, R., Dionísio, A., & Hassani, H. (2012). On the globalization of stock markets: an application of vector error correction model, mutual information and singular spectrum analysis to the G7 countries. *The Quarterly Review of Economics and Finance*, *52*(4), 369-384. DOI: https://doi.org/10.1016/j.qref.2012.10.002

Mizuno, T., Kurihara, S., Takayasu, M., & Takayasu, H. (2004). Time-scale dependence of correlations among foreign currencies. In *The Application of Econophysics* (pp. 24-29). Springer, Tokyo. DOI: http://doi.org/fzwb9v

Morse, M., & Hedlund, G. A. (1938). Symbolic dynamics. *American Journal of Mathematics*, *60*(4), 815-866. DOI: http://dx.doi.org/10.2307/2371264

Moselhy, H. F., Georgiou, G., & Kahn, A. (2001). Frontal lobe changes in alcoholism: a review of the literature. *Alcohol and alcoholism*, *36*(5), 357-368. DOI: https://doi.org/10.1093/alcalc/36.5.357

Murphy, G. I. (1964). *Species replacement in marine ecosystems with reference to the California current*. Scripps Institution of Oceanography. DOI: https://doi.org/10.1525/j.ctv1xxzp6.

Musmeci, N., Aste, T., & Di Matteo, T. (2016). Interplay between past market correlation structure changes and future volatility outbursts. *Scientific reports*, *6*, 36320. DOI: http://dx.doi.org/10.1038/srep36320

Musmeci, N., Aste, T., & Di Matteo, T. (2014). Risk diversification: a study of persistence with a filtered correlation-network approach. *arXiv preprint arXiv:1410.5621*. DOI: https://doi.org/10.1410.5621v1

Musmeci, N., Nicosia, V., Aste, T., Di Matteo, T., & Latora, V. (2017). The multiplex dependency structure of financial markets. *Complexity*, *2017*. DOI: https://doi.org/10.1155/2017/9586064

Neutel, A. M., Heesterbeek, J. A., Van de Koppel, J., Hoenderboom, G., Vos, A., Kaldeway, C., ... & De Ruiter, P. C. (2007). Reconciling complexity with stability in naturally assembling food webs. *Nature*, *449*(7162), 599. DOI: http://dx.doi.org/10.1038/nature06154

Newman, M. E. (2002). Assortative mixing in networks. *Physical review letters*, *89*(20), 208701. DOI: https://doi.org/10.1103/PhysRevLett.89.208701

Newman, M. E. (2016). Mathematics of networks. *The new Palgrave dictionary of economics*, 1-8. DOI: https://doi.org/10.1108/09504120310503827

Packard, N. H., Crutchfield, J. P., Farmer, J. D., & Shaw, R. S. (1980). Geometry from a time series. *Physical review letters*, *45*(9), 712. DOI: https://doi.org/10.1103/PhysRevLett.45.712

Patil, D. A. S., Hunt, B. R., & Carton, J. A. (2001). Identifying low-dimensional nonlinear behavior in atmospheric data. *Monthly weather review*, *129*(8), 2116-2125. DOI: https://doi.org/10.1175/1520-0493(2001)129<2116:ILDNBI>2.0.CO;2

Pearl, J. (2003). Causality: Models, Reasoning, and Inference. Cambridge University Pres; 2000. DOI: http://dx.doi.org/10+10170S0266466603004109

Péguin-Feissolle, A., Strikholm, B., & Teräsvirta, T. (2013). Testing the Granger noncausality hypothesis in stationary nonlinear models of unknown functional form. *Communications in Statistics-Simulation and Computation*, *42*(5), 1063-1087. DOI: https://doi.org/10.1080/03610918.2012.661500

Pearson, K. (1895). VII. Note on regression and inheritance in the case of two parents. *proceedings of the royal society of London*, *58*(347-352), 240-242. DOI: https://doi.org/10.1098/rspl.1895.0041

Peltonen, T. A., Scheicher, M., & Vuillemey, G. (2014). The network structure of the CDS market and its determinants. *Journal of Financial Stability*, *13*, 118-133. DOI: https://doi.org/10.1016/j.jfs.2014.05.004

Pfefferbaum, A., Sullivan, E. V., Mathalon, D. H., & Lim, K. O. (1997). Frontal lobe volume loss observed with magnetic resonance imaging in older chronic alcoholics. *Alcoholism: Clinical and Experimental Research*, *21*(3), 521-529. DOI: https://doi.org/10.1111/j.1530-0277.1997.tb03798.x

Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335-346. DOI: https://doi.org/10.1093/biomet/75.2.335

Pijn, J. P. M., Vijn, P. C. M., Lopes da Silva, F. H., Van Emde Boas, W., & Blanes, W. (1989). The use of signal-analysis for the localization of an epileptogenic focus: a new approach. *Adv Epileptol*, *17*, 272-276. DOI: https://doi.org/17:272-276

Pijn, J. P. M., Vijn, P. C. M., Da Silva, F. L., Boas, W. V. E., & Blanes, W. (1990). Localization of epileptogenic foci using a new signal analytical approach. *Neurophysiologie Clinique/Clinical Neurophysiology*, *20*(1), 1-11. DOI: https://doi.org/10.1016/S0987-7053(05)80165-0

Pons, P., & Latapy, M. (2005, October). Computing communities in large networks using random walks. In *International symposium on computer and information sciences* (pp. 284-293). Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/11569596_31

Porter, M. A., Onnela, J. P., & Mucha, P. J. (2009). Communities in networks. *Notices of the AMS*, *56*(9), 1082-1097. DOI: https://doi.org/10.0902.3788

Prechter Jr, R. R. (2001). Unconscious herding behavior as the psychological basis of financial market trends and patterns. *The Journal of Psychology and Financial Markets*, *2*(3), 120-125. DOI: https://doi.org/10.1207/S15327760JPFM0203_1

Preis, T., Reith, D., & Stanley, H. E. (2010). Complex dynamics of our economic life on different scales: insights from search engine query data. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *368*(1933), 5707-5719. DOI: https://doi.org/10.1098/rsta.2010.0284

Preis, T., Schneider, J. J., & Stanley, H. E. (2011). Switching processes in financial markets. *Proceedings of the National Academy of Sciences*, *108*(19), 7674-7678. DOI: https://doi.org/10.1073/pnas.1019484108

Qiao, Z., & Lam, K. S. (2011). Granger causal relations among Greater China stock markets: a nonlinear perspective. *Applied Financial Economics*, *21*(19), 1437-1450. DOI: https://doi.org/10.1080/09603107.2011.577007

Rahman, M., & Mustafa, M. (1997). Dynamic linkages and Granger causality between short-term US corporate bond and stock markets. *Applied Economics Letters*, *4*(2), 89-91. DOI: https://doi.org/10.1080/758526701

Ratti, M. T., Bo, P., Giardini, A., & Soragna, D. (2002). Chronic alcoholism and the frontal lobe: which executive functions are imparied?. *Acta Neurologica Scandinavica*, *105*(4), 276-281. DOI: https://doi.org/10.1034/j.1600-0404.2002.0o315.x

Raven, P. H., & Johnson, G. B. (1999). Biology, Boston: Hill Companies. DOI: https://doi.org/10.978-1259188138

Russo, J. E., Schoemaker, P. J., & Russo, E. J. (1989). *Decision traps: Ten barriers to brilliant decision-making and how to overcome them*. New York, NY: Doubleday/Currency. DOI: https://doi.org/10.1002.3960040106

Sandoval, L. (2014). Structure of a global network of financial companies based on transfer entropy. *Entropy*, *16*(8), 4443-4482. DOI: https://doi.org/10.3390/e16084443

Sauer, T., Yorke, J. A., & Casdagli, M. (1991). Embedology. *Journal of statistical Physics*, *65*(3-4), 579-616. DOI: https://doi.org/10.1007/bf01053745

Scheinkman, J. A., & LeBaron, B. (1986). Nonlinear dynamics and stock returns.

Schreiber, T. (2000). Measuring information transfer. *Physical review letters*, *85*(2), 461. DOI: https://doi.org/10.1103/PhysRevLett.85.461

Seidman, S. B. (1983). Network structure and minimum degree. *Social networks*, *5*(3), 269-287. DOI: https://doi.org/10.1016/0378-8733(83)90028-X

Sensoy, A., Sobaci, C., Sensoy, S., & Alali, F. (2014). Effective transfer entropy approach to information flow between exchange rates and stock markets. *Chaos, solitons & fractals*, *68*, 180-185. DOI: https://doi.org/10.1016/j.chaos.2014.08.007

Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, *27*(3), 379-423. DOI: ht tps://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Siliverstovs, B., L'Hégaret, G., Neumann, A., & Von Hirschhausen, C. (2005). International market integration for natural gas? A cointegration analysis of prices in Europe, North America and Japan. *Energy Economics*, *27*(4), 603-615. DOI: https://doi.org/10.1016/j.eneco.2005.03.002

Simpson, E. S. W., & Fedorov, K. N. (1983). Proceedings of the joint oceanographic assembly 1982 general symposium, Dalhousie University, Halifax, Nova Scotia, Canada. DOI: https://scor-int.org/Publications/1982Proceedings.PDF

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology: Human learning and memory*, *6*(2), 174. DOI: http://dx.doi.org/10.1037/0278-7393.6.2.174

Stambaugh, R.F., 2014. Presidential address: Investment noise and trends. The Journal of Finance, 69(4), pp.1415-1453. https://doi.org/10.1111/jofi.12174

Stavroglou, S. K., Pantelous, A. A., Soramaki, K., & Zuev, K. (2017). Causality networks of financial assets. *Journal of Network Theory in Finance*, *3*(2), 17-67. DOI: https://doi.org/10.21314/JNTF.2017.029

Stavroglou, S. K., Pantelous, A. A., Stanley, H. E., & Zuev, K. M. (2019). Hidden interactions in financial markets. *Proceedings of the National Academy of Sciences*, *116*(22), 10646-10651. DOI: https://doi.org/10.1073/pnas.1819449116

Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *science*, *338*(6106), 496-500. DOI: https://doi.org/10.1126/science.1227079

Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980* (pp. 366-381). Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/BFb0091924

Tett, G. (2006). The dream machine: invention of credit derivatives. *Financial Times*, *24*. https://www.ft.com/content/7886e2a8-b967-11da-9d02-0000779e2340

Toulmin, S. E. (1961). Foresight and understanding: An enquiry into the aims of science. DOI: https://books.google.co.ke/books/about/Foresight_and_Understanding.html?id=P24gAQAAMAAJ&redir_esc=y

Trimble, V. (1987). Existence and nature of dark matter in the universe. *Annual review of astronomy and astrophysics*, *25*(1), 425-472. DOI: https://doi.org/10.1146/annurev.aa.25.090187.002233

Tumminello, M., Aste, T., Di Matteo, T., & Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences*, *102*(30), 10421-10426. DOI: https://doi.org/10.1073/pnas.0500298102

Valone, T. J., & Balaban-Feld, J. (2019). An experimental investigation of top–down effects of consumer diversity on producer temporal stability. *Journal of Ecology*, *107*(2), 806-813. DOI: https://doi.org/10.1111/1365-2745.13064

Valone, T. J., & Balaban-Feld, J. (2018). Impact of exotic invasion on the temporal stability of natural annual plant communities. *Oikos*, *127*(1), 56-62. DOI: https://doi.org/10.1111/oik.04591

Výrost, T., Lyócsa, Š., & Baumöhl, E. (2015). Granger causality stock market networks: Temporal proximity and preferential attachment. *Physica A: Statistical Mechanics and its Applications*, *427*, 262-276. DOI: https://doi.org/10.1016/j.physa.2015.02.017

Wang, D., Podobnik, B., Horvatić, D., & Stanley, H. E. (2011). Quantifying and modeling long-range cross correlations in multiple time series with applications to world stock indices. *Physical Review E*, *83*(4), 046121. DOI: https://doi.org/10.1103/PhysRevE.83.046121

Wang, Z., Yang, J., & Li, Q. (2007). Interest rate linkages in the Eurocurrency market: Contemporaneous and out-of-sample Granger causality tests. *Journal of International Money and Finance*, *26*(1), 86-103. DOI: https://doi.org/10.1103/PhysRevE.83.046121https://doi.org/10.1016/j.jimonfin.2006.10.005

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *nature*, *393*(6684), 440. DOI: https://doi.org/10.1038/30918

Weigend, A. S. (2018). *Time series prediction: forecasting the future and understanding the past*. Routledge. DOI: https://doi.org/10.4324/9780429492648

Wendling, F., Bartolomei, F., Bellanger, J. J., & Chauvel, P. (2001). Interpretation of interdependencies in epileptic signals using a macroscopic physiological model of the EEG. *Clinical neurophysiology*, *112*(7), 1201-1218. DOI: https://doi.org/10.1016/S1388-2457(01)00547-8

West, D. B. (1996). *Introduction to graph theory* (Vol. 2). Upper Saddle River, NJ: Prentice hall. http://free-journal.umm.ac.id

Whitney, H. (1936). Differentiable manifolds. *Annals of Mathematics*, 645-680. DOI: https://doi.org/10.2307/1968482

Yang, C., Chen, Y., Niu, L., & Li, Q. (2014). Cointegration analysis and influence rank—A network approach to global stock markets. *Physica A: Statistical Mechanics and its Applications*, *400*, 168-185. DOI: https://doi.org/10.1016/j.physa.2014.01.011

Yook, S. H., Chae, H., Kim, J., & Kim, Y. (2016). Finding modules and hierarchy in weighted financial network using transfer entropy. *Physica A: Statistical Mechanics and its Applications*, *447*, 493-501. DOI: https://doi.org/10.1016/j.physa.2015.12.018

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between Time-Series?--a study in sampling and the nature of time-series. *Journal of the royal statistical society*, *89*(1), 1-63. DOI: https://doi.org/10.2307/2341482

Zhang, X. L., Begleiter, H., Porjesz, B., Wang, W., & Litke, A. (1995). Event related potentials during object recognition tasks. *Brain Research Bulletin*, *38*(6), 531-538. DOI: https://doi.org/10.1016/0361-9230(95)02023-5

Zhang, Y. J., & Wei, Y. M. (2010). The crude oil market and the gold market: Evidence for cointegration, causality and price discovery. *Resources Policy*, *35*(3), 168-177. DOI: https://doi.org/10.1016/j.resourpol.2010.05.003

Zhou, P., Lu, F., & Wang, S. (2014). Testing linear and nonlinear granger causality in CSI300 futures and spot markets based on new concepts of nonlinear positive/negative spillover. *Journal of Systems Science and Complexity*, *27*(4), 729-742. DOI: https://doi.org/10.1007/s11424-014-2261-3

Zhou, S., & Mondragón, R. J. (2004). The rich-club phenomenon in the Internet topology. *IEEE Communications Letters*, *8*(3), 180-182. DOI: https://doi.org/10.1109/LCOMM.2004.823426